

7MNB

*Journal of*  
**Geophysical  
Research**

**VOLUME 65 . SEPTEMBER 1960 NUMBER 9**

**THE SCIENTIFIC PUBLICATION  
OF THE AMERICAN GEOPHYSICAL UNION**

# Journal of Geophysical Research

*An International Scientific Publication*

## OFFICERS OF THE UNION

LLOYD V. BERKNER, *President*  
THOMAS F. MALONE, *Vice President*  
A. NELSON SAYRE, *General Secretary*  
WALDO E. SMITH, *Executive Secretary*

## OFFICERS OF THE SECTIONS

### Geodesy

CHARLES PIERCE, *President*  
FLOYD W. HOUGH, *Vice President*  
BUFORD K. MEADE, *Secretary*

### Seismology

LEONARD M. MURPHY, *President*  
JAMES A. PEOPLES, JR., *Vice President*  
BENJAMIN F. HOWELL, JR., *Secretary*

### Meteorology

THOMAS F. MALONE, *President*  
GORDON E. DUNN, *Vice President*  
WOODROW C. JACOBS, *Secretary*

### Geomagnetism and Aeronomy

L. R. ALDREDGE, *President*  
C. T. ELVEY, *Vice President*  
J. HUGH NELSON, *Secretary*

### Oceanography

WALTER H. MUNK, *President*  
DONALD W. PRITCHARD, *Vice President*  
EUGENE C. LAFOND, *Secretary*

### Volcanology, Geochemistry, and Petrology

ALFRED O. C. NIER, *President*  
FRANCIS J. TURNER, *Vice President*  
IRVING FRIEDMAN, *Secretary*

### Hydrology

WALTER B. LANGBEIN, *President*  
WILLIAM C. ACKERMANN, *Vice President*  
CHARLES C. McDONALD, *Secretary*

### Tectonophysics

PATRICK M. HURLEY, *President*  
LOUIS B. SLICHTER, *Vice President*  
H. RICHARD GAULT, *Secretary*

## BOARD OF EDITORS

Editors: PHILIP H. ABELSON and J. A. PEOPLES

## ASSOCIATE EDITORS

1960

|                 |                    |
|-----------------|--------------------|
| HENRY G. BOOKER | WALTER B. LANGBEIN |
| E. C. BULLARD   | ERWIN SCHMID       |
| JULE CHARNEY    | HENRY STOMMEL      |
| GEORGE T. FAUST | J. TH. THIJSSSE    |
| DAVID G. KNAPP  | A. H. WAYNICK      |

J. TUZO WILSON

1960-1961

|                |                 |
|----------------|-----------------|
| HENRI BADER    | T. NAGATA       |
| K. E. BULLEN   | FRANK PRESS     |
| CONRAD P. MOOK | A. NELSON SAYRE |
| WALTER H. MUNK | MERLE A. TUVE   |

JAMES A. VAN ALLEN

1960-1962

|                        |                 |
|------------------------|-----------------|
| JULIUS BARTELS         | L. A. MANNING   |
| V. V. BELOUSSON        | TOR J. NORDENSC |
| E. G. BOWEN            | E. N. PARKER    |
| JOHN E. CHAPPELEAR     | GEORGE P. RIGSB |
| G. D. GARLAND          | WALTER O. ROBE  |
| GORDON J. F. MACDONALD | C. N. TOUART    |

JAMES R. WAIT

This Journal welcomes original scientific contributions on the physics of the earth and its environment.

Manuscripts should be transmitted to J. A. Peopes, Jr., Geology Department, University of Kansas, Lawrence, Kansas. Authors' institutions, if in the United States or Canada, are requested to pay a publication charge of \$25 per page, which, if honored, entitles them to 100 free reprints.

Subscriptions to the *Journal of Geophysical Research* and *Transactions, AGU* are included in membership dues.

Non-member subscriptions, *Journal of Geophysical Research*...\$30 for back Volume of 1959, \$15 for this issue; \$20 for the calendar year 1960.

Non-member subscriptions, *Transactions, AGU*.....\$4 per calendar year, \$1.25 per page.

Subscriptions, renewals, and orders for back numbers should be addressed to American Geophysical Union, 1515 Massachusetts Ave., Northwest, Washington, D. C. Suggestions to authors are available on request.

Advertising Representative: Howland and Howland, Inc., 230 Park Ave., New York 17, N. Y.

Beginning with the January 1959 issue (Vol. 64, No. 1) the *Journal of Geophysical Research* is published monthly by the American Geophysical Union, the U. S. National Committee of the International Union of Geodesy and Geophysics organized under the National Academy of Sciences-National Research Council as the U. S. national adhering body. Publication of this journal is supported by the National Science Foundation and the Carnegie Institution of Washington. The new monthly combines the type of scientific material formerly published in the bi-monthly *Transactions, American Geophysical Union*, and the quarterly *Journal of Geophysical Research*. The *Transactions, American Geophysical Union* will continue as a quarterly publication for Union business and items of interest to members of the Union.

Published monthly by the American Geophysical Union from 1407 Sherwood Avenue, Richmond, Virginia. Second class postage paid at Richmond, Virginia.



# STRANGE "FISH" UNDER THE POLAR ICE

## Revolutionary RCA Magnetic Video Tape Recorder to Speed Navigation Training of Submariners

Aboard the nuclear submarine Sea Dragon, the first undersea magnetic video tape recorder will record and store data on under-the-ice characteristics from externally installed TV cameras. Upon return to base the recorded information will be displayed for the benefit of undersea service trainees.

The RCA undersea recorder is a marvel of compact design (dimensions 20" x 20" x 100"). It nestles in a torpedo rack, and represents a 60% space reduction over existing video tape equipment.

Among the exclusive RCA developments are: the now famous "Tiros" satellite recorder; a radar system designed to take the first pictures of a nose cone re-entry vehicle; a unique tape cartridge completely adaptable to any size recorder. For literature describing new RCA defense and commercial products developments, write Defense Electronic Products, Radio Corporation of America, Camden, N. J.

Out of today's defense needs... tomorrow's electronic advances



The Most Trusted Name  
in Electronics

RADIO CORPORATION OF AMERICA





# "THE NOBLE SCIENCE OF GEOLOGY LOSES GLORY FROM THE EXTREME IMPERFECTION OF THE RECORD"

*Darwin*

Since the days of Darwin "the noble science of geology" has gained much in glory. Meanwhile an understanding of the environment in which men live has continued to be necessary to his survival. Yet, today, as man seeks to expand from his earth into space, he finds that his knowledge of the earth is in many senses inadequate to the job. Our knowledge of the oceans, of the roots of the mountains, and of the nature of the earth's crust has become, quite suddenly, closely connected to our ability to survive.

The exploration geophysics industry can supply basic scientific methods, necessary techniques, and many of the skilled, experienced personnel to solve these problems in a manner similar to the ways in which this industry has discovered several hundred billion barrels of petroleum in the last twenty-five years. This industry is confident of its ability to contribute significantly to the solution of problems in the Earth Sciences that the Free World faces today.

*General*

**GEOPHYSICAL COMPANY**

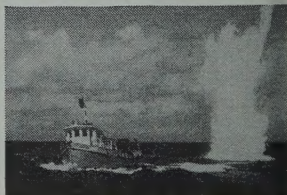
HOUSTON CLUB BUILDING • HOUSTON, TEXAS

In Canada: 10509 81st Avenue,  
Edmonton, Alberta, Canada

General Geophysical Company de France (SARL),  
4 Square Rapp • Paris 7, France

General Geophysical Company (Bahamas) Ltd., Libyan Branch  
22, Via Vittorio Gadolini, Tripoli Libya  
General Geophysical Company (Bahamas) Ltd.  
Umm Said, Quatar, Arabian Gulf

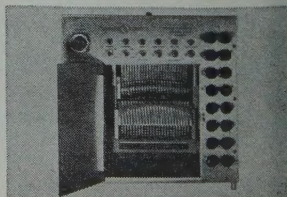
The General Geophysical Company has been a leader in the geophysical industry since 1935. With its large laboratory and over 600 employees with worldwide geophysical experience, General Geophysical Company is prepared to develop special instruments, operate field stations, and interpret geophysical data on many aspects of studies of the earth's crust and the oceans.



**EXPERIENCE**



**PERSONNEL**

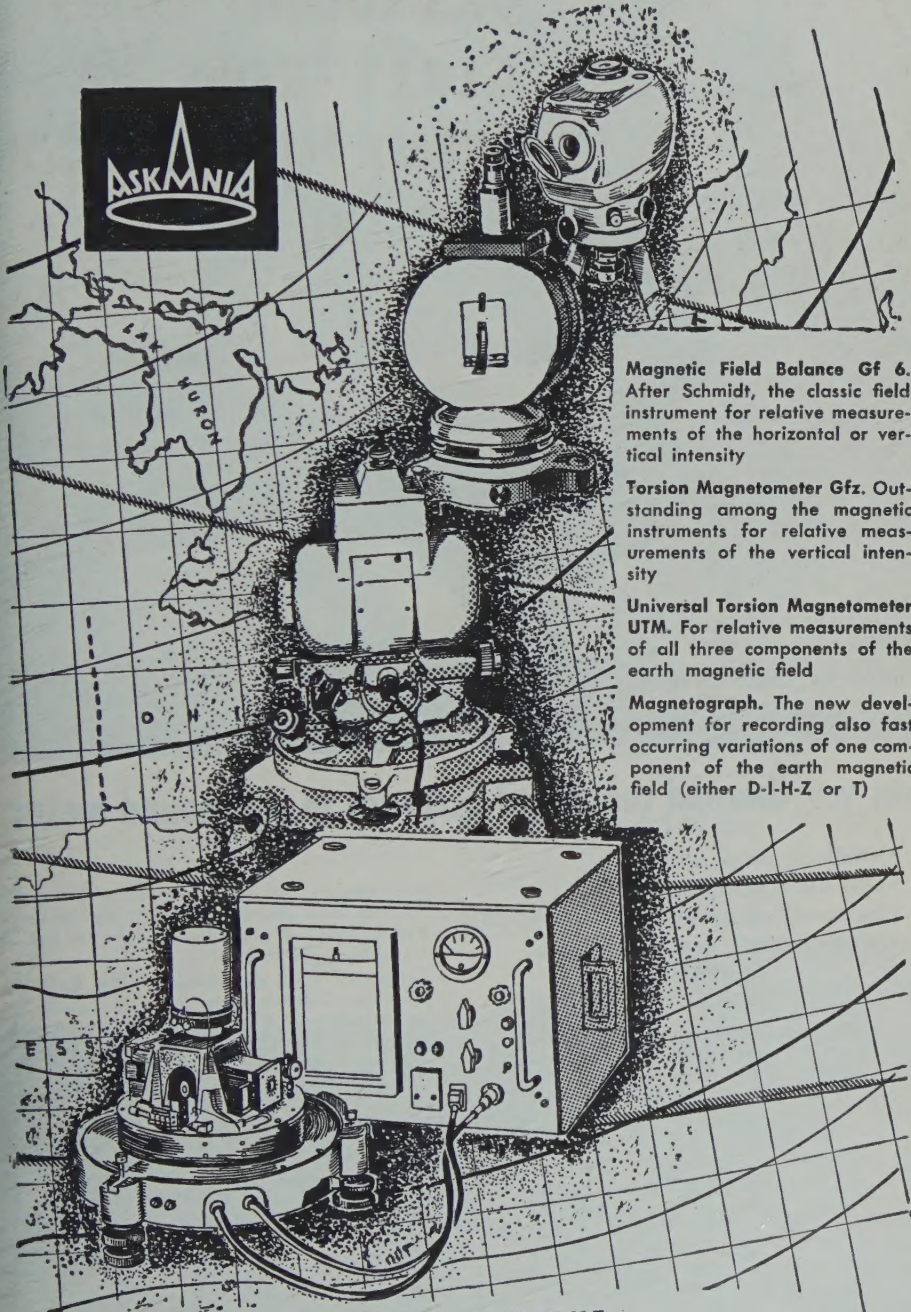


**EQUIPMENT**

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers



# ASKANIA Magnetic Instruments for precise determination of magnetic anomalies



**Magnetic Field Balance Gf 6.** After Schmidt, the classic field instrument for relative measurements of the horizontal or vertical intensity

**Torsion Magnetometer Gfz.** Outstanding among the magnetic instruments for relative measurements of the vertical intensity

**Universal Torsion Magnetometer UTM.** For relative measurements of all three components of the earth magnetic field

**Magnetograph.** The new development for recording also fast occurring variations of one component of the earth magnetic field (either D-I-H-Z or T)

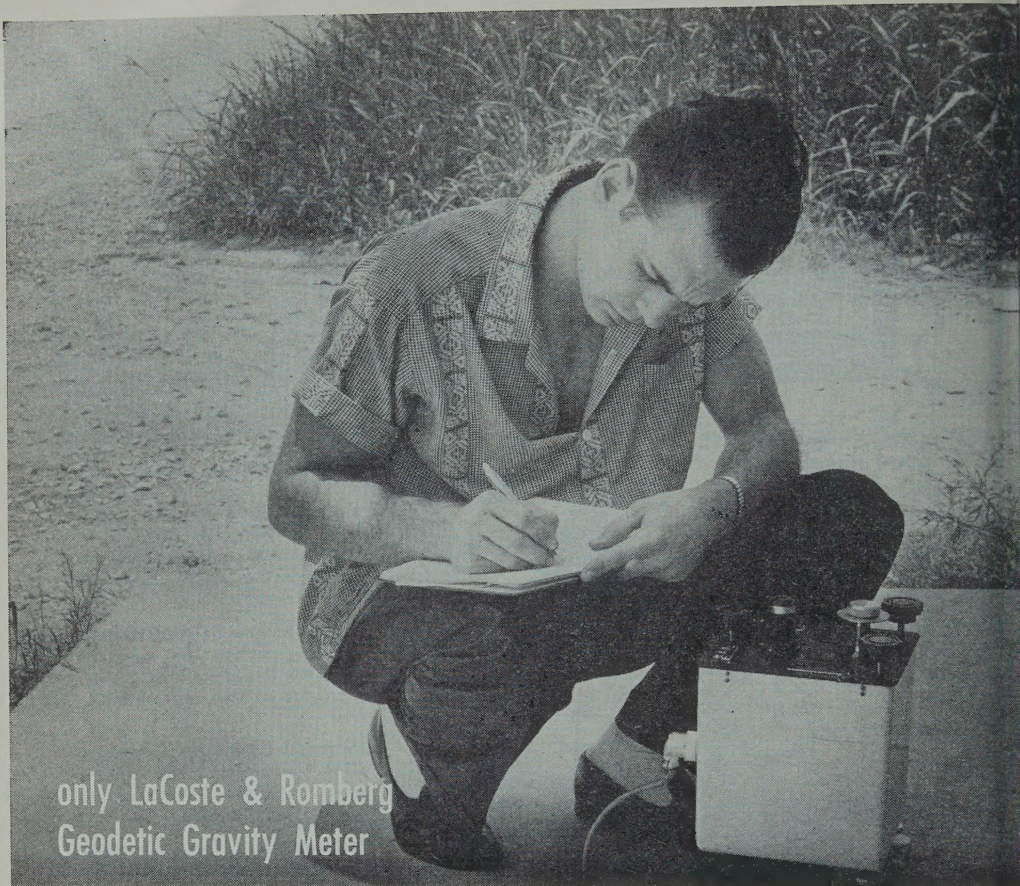
**ASKANIA-WERKE**

U. S. Branch Office & Service Dept.

4913 Cordell Ave., Bethesda, Maryland

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers





only LaCoste & Romberg  
Geodetic Gravity Meter

**gives you thermal controlled accuracy in a 7-pound meter**

- world survey without resetting
- never requires recalibrating
- less than 0.5 mgl drift per month
- no "sets" or "tares" under normal operation

This new miniaturized Geodetic Gravity Meter retains all the accuracy and dependability of the standard model introduced by LaCoste & Romberg in 1956, yet it weighs only 7 pounds. (Complete with battery and luggage-type carrying case, it weighs less than 17 pounds). With a world-wide range of over 6,000 mgl., this instrument has a repeatability of 0.01 mgl. Actual field tests over the complete gravity range have shown an accuracy better than 0.04 mgl.

Exceptionally high sensitivity of the LaCoste & Romberg meter is attained by a zero length spring suspension (U. S. Patent No. 2,293,437). Calibration is stabilized by means of patented

lever systems that act on the main spring rather than on weak measuring springs. And by thermostating, drift is normally reduced to less than 0.5 milligal per month.

Rugged and dependable, the LaCoste & Romberg Geodetic Meter requires practically no maintenance in the field. Its gravity responsive system is completely suspended by springs and will therefore withstand any shock that will not damage the housing supporting it. It is specifically designed to provide a light-weight meter with higher accuracy and lower drift than can be attained in any other geodetic gravity meter. For complete information, write for *Miniature Geodetic Gravity Meter Bulletin*.



**LaCoste & Romberg**

6606 NORTH LAMAR

AUSTIN, TEXAS

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertiser



# geophysical research

at Hughes Research Laboratories

Basic and applied studies at Hughes Research Laboratories include many and varied geophysical research programs. Among these are:

**Upper Atmosphere & Space Physics**—basic studies in the plasma physics of the atmosphere, solar-terrestrial relationships, properties of geomagnetically trapped radiation, and analysis of artificially produced ionization effects.

**Radio-wave Propagation**—radar echoes from auroral and other natural ionization, scatter propagation, fading and diversity reception theory, VLF and whistler mode propagation, and microwave and infrared transmission characteristics.

**Physical Oceanography**—long-range acoustic propagation, specialized phenomena of importance to submarine detection and communications, and commercial applications of oceanography.

If your research interests are related to the above fields, your inquiry is encouraged. The new facilities of the Hughes Research Laboratories are located in Malibu, California—overlooking the Pacific Ocean. Here scientists and engineers are offered the academic atmosphere which is essential for effective research effort.

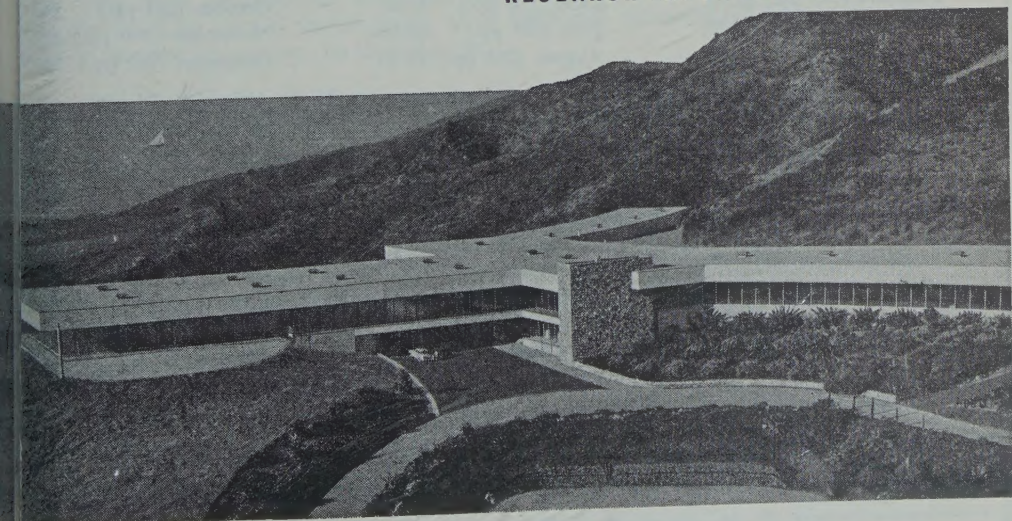
*Your inquiry regarding staff openings may be directed in confidence to: Mr. D. A. Bowdoin, HUGHES RESEARCH LABORATORIES, Malibu 10, California.*

*Creating a new world with ELECTRONICS*

**HUGHES**

HUGHES AIRCRAFT COMPANY

**RESEARCH LABORATORIES**



Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers



# SPRENGNETHER LONG PERIOD HORIZONTAL SEISMOMETER

## GENERAL SPECIFICATIONS:

- **Period Range:** 6 to 70 seconds.
- **Magnification:** Up to 15,000, depending on operating period.
- **Damping:** Electromagnetic.
- **Transducer-moving coils in circular magnetic gaps.**
- **Coils approximately 500 ohms. or to your specifications.**
- **Pendulum steady mass weight 22 lbs.**

## PHYSICAL SPECIFICATIONS:

Length.....25"  
Width.....25"  
Height.....22"  
Net Weight.....90 lbs. approx.  
Shipping Weight.....155 lbs. approx.

## COVER:

Light weight aluminum with rubber gasket seal to base.  
Provided with viewing window to observe pendulum centering.

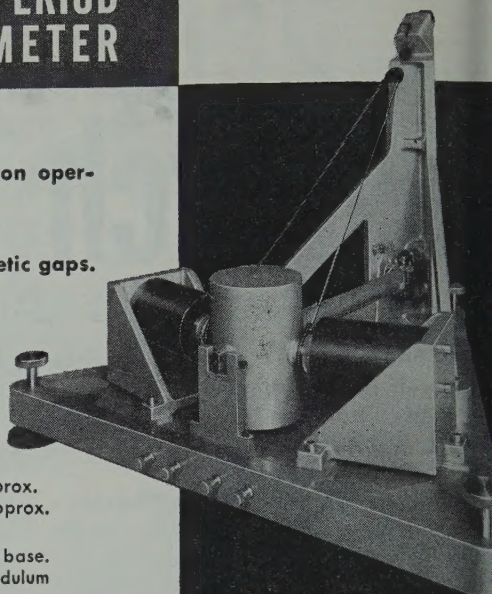
*To compliment this instrument,  
a long period vertical seismometer is also available.*

Internationally Known Mfrs. of Seismological, Geophysical Instruments.

**W. F. SPRENGNETHER INSTRUMENT CO., INC.**

4567 SWAN AVENUE

ST. LOUIS 10, MO.



**WRITE FOR DETAIL  
INFORMATION ON THE  
INSTRUMENTS.**

## Back issues available . . .

### Journal of Geophysical Research • 1959 • Volume 64

Complete set of 12 (2488 pp.)—\$30.00

|                           |                         |                            |
|---------------------------|-------------------------|----------------------------|
| January (132 pp.) \$2.00  | May (98 pp.) \$2.00     | September (230 pp.) \$3.00 |
| February (138 pp.) \$2.00 | June (110 pp.) \$2.00   | October (284 pp.) \$4.00   |
| March (112 pp.) \$2.00    | July (168 pp.) \$2.00   | November (390 pp.) \$5.00  |
| April (106 pp.) \$2.00    | August (268 pp.) \$4.00 | December (452 pp.) \$6.00  |

Scientific Effects of Artificially Introduced Radiations at High Altitudes (74-page Symposium reprinted from the August issue) ..... \$1.50

International Symposium on Fluid Mechanics in the Ionosphere (202-page Symposium reprinted from the December issue) ..... \$4.50

Subscriptions to Journal of Geophysical Research, Calendar Year 1960..... \$20.00

## AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Avenue, N.W. • Washington 5, D. C.

*Note:* The former publisher of the *Journal of Geophysical Research* has turned over the stock of back numbers, 1958 and earlier, to Walter J. Johnson, Inc., 111 Fifth Avenue, New York 3, New York.

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertiser



**FOUR IMPORTANT NEW BOOKS  
FROM MCGRAW-HILL**

**IGNEOUS AND METAMORPHIC PETROLOGY, New Second Edition**

by FRANCIS J. TURNER and JOHN VERHOOGEN, *both of the University of California, Berkeley*. 694 pages, \$12.00

As before, the book represents a unified general impression of origin and evolution of rocks that have crystallized, or have been profoundly modified, at high temperatures. It is correlated with modern conceptions as to the nature and prevailing physical conditions of the earth's crust and of the outer part of the underlying mantle. Igneous and metamorphic phenomena have been treated, in a single volume, as partially dependent on each other, and as being controlled by the same general physico-chemical principles.

**INTRODUCTION TO GEOPHYSICAL PROSPECTING, New Second Edition**

by MILTON B. DOBRIN, *Triad Oil Company, Ltd. Calgary, Alberta, Canada*. 446 pages, \$9.50.

A thorough revision of a highly successful text. It is designed to present the principles of current techniques of geophysical prospecting for oil and minerals to students and technical personnel employed in the fields of petroleum and mineral exploration. The book covers all the major methods of geophysical prospecting. For each method it discusses fundamental physical principles, instruments, field techniques, reduction of data, interpretation, and examples showing results of actual surveys.

**PHYSICAL GEOLOGY: Principles and Processes, New Fifth Edition**

by WILLIAM H. EMMONS, IRA S. ALLISON, *Oregon State College*; GEORGE A. THIEL, *University of Minnesota*; and CLINTON R. STAUFFER, *California Institute of Technology*. 491 pages, \$7.95.

A textbook for a beginning one-semester course in Physical Geology. It has been a standard and leading text for this field for over 27 years. This new fifth edition has been the most extensive of all its revisions. It is receiving special editing treatment and the result will be a completely rewritten text, with entirely new artwork done by a professional scientific illustrator, and a skilled use of two colors throughout. The text is suitable for both terminal cultural courses, or for introductory courses for majors in the subject.

**PRINCIPLES OF PETROLEUM GEOLOGY, New Second Edition**

by WILLIAM L. RUSSELL, *Agricultural and Mechanical College of Texas*. 503 pages, \$9.50.

This new edition of one of the most important books in this area has been carefully revised; much new material has been added, and all other information has been brought fully up to date. The first 17 chapters deal with principles; the remaining 12 chapters cover the specialized methods or techniques used by petroleum geologists. The book will be of value, therefore, not only to students, but also to those involved in the oil industry.

Send for Copies on Approval

**McGRAW-HILL BOOK COMPANY, INC.**

330 West 42nd Street

New York 36, N. Y.

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers



## NEW BOOKS FROM INTERSCIENCE

### Methods and Techniques in Geophysics

edited by S. K. RUNCORN, *King's College, University of Durham*

Modern advances in physical techniques in this field are presented here by an international group of experts to help fill a relative void in the literature of pure geophysical research. The topics discussed in this first volume include *Measurement of: Temperature Gradient in the Earth, Heat Flow over Land, Gravity at Sea, The Geomagnetic Elements; also Bore-Hole Surveying, Measurements in Paleomagnetism, The Detection of Earth Movements, Earth Currents, Properties of Rocks under High Pressure and Temperature, Latitude and Longitude, and the Secular Motion of the Pole.* An additional volume is in preparation.

Volume 1 387 pages \$10.00

### Lead Isotopes in Geology

by R. D. RUSSELL, *University of British Columbia*, and R. M. FARQUHAR, *University of Toronto*

The application of isotopic studies to geophysics has developed rapidly in very recent years, and there is need for this extensive detailed description. The authors, who have contributed much to rock dating by lead isotope determinations, have surveyed the field and given a coherent and up-to-date account of the various ideas that they and other investigators have expressed on the subject. A feature is the very extensive tabulation of lead isotopic abundances, from the authors' own work in Toronto and from other centers of geophysical studies in many parts of the world.

251 pages \$9.00

### Methods and Techniques in Geochemistry

edited by A. A. SMALES, *United Kingdom Atomic Energy Research Establishment, Harwell*, and L. R. WAGER, *Oxford*

In this single volume, experts consider the major analytical methods used in geochemistry. Recent advances are discussed, as well as older techniques and their new developments. Each major method, with various related techniques, is treated in a separate contributed chapter, including a discussion of the underlying theory. Thoroughly practical, an early chapter deals with the preliminary examination and treatment of the material before analysis begins.

471 pages \$13.50

**INTERSCIENCE PUBLISHERS, INC.**

250 Fifth Ave., New York 1, N. Y.

## ARE YOU A MEMBER of AGU?

If you are a reader or user of either the publications noted below, you should look into this matter.

Members regularly receive  
the monthly

**Journal of  
Geophysical  
Research**

the quarterly

**Transactions  
American  
Geophysical  
Union**

Members are entitled to special counts on other publications of AGU such as

- **Geophysical Monograph Series** (Number 1 to 5 available)
- **Izvestiya of the Academy of Sciences USSR; Geophysics Series** for the years 1957, 1958, and 1959

- 
- **Annual meeting in Washington** in April or early May each year
  - **Regional meeting in the Pacific Northwest** in autumn
  - **Regional meeting in the Pacific Southwest** in winter
  - Others as opportunity permits

For application forms and other information regarding membership, meetings, and other matters, write to

**AMERICAN GEOPHYSICAL UNION**

**1515 Massachusetts Ave., N.W.  
Washington 5, D. C.**

(An application form is also inserted in the closing pages of this issue.)

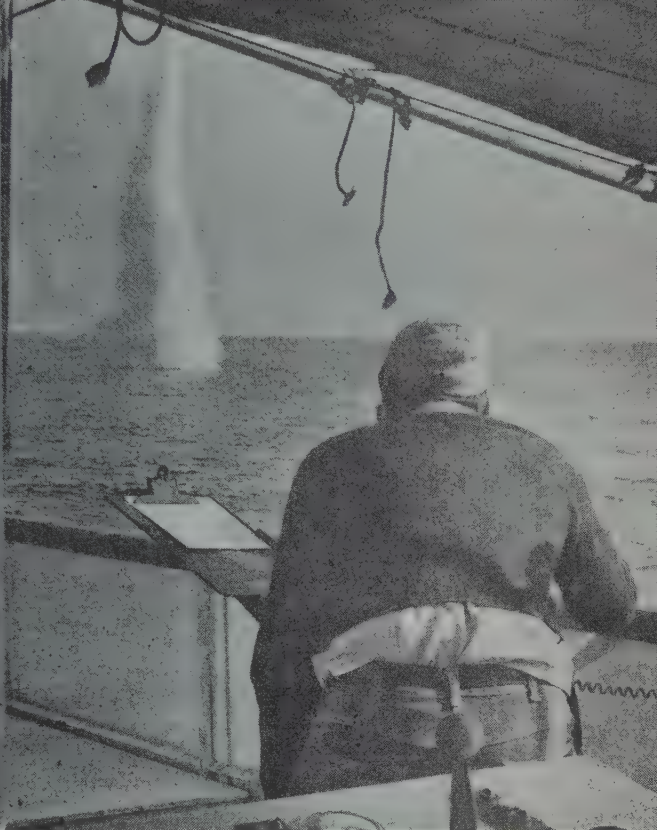




### TO THE OCEAN'S FLOOR ... AND BELOW

Daily aboard TI's 405-ton, 158-foot SONIC—uniquely a complete sea-going seismic and acoustic investigation center—geoscientists are studying and mapping the structure of the earth below the ocean's floor. Illustrated above is the seismic reflection method, with explosive charges detonated nearby for detailed study of energy travel through water and rock. Technicians aboard the SONIC also record refracted seismic energy, measuring signals through the ocean.

Explosive charges exploded on land or in water as much as 35 miles away. Similar land-based TI geophysical parties are now operating in more than 22 countries of the world.



INNER-SPACE KNOWLEDGE FOR OUTER-SPACE TECHNOLOGY ADVANCED BY TI'S

## scientific key to Davy Jones' locker!

At a time when the focus is on outer space—a time when we know more about the moon's surface than we know about 70% of the earth's surface—geophysicists such as this sea-going scientist from Texas Instruments are furthering space-age technology with studies of inner space of the oceans.

TI represents one of more than 60 global TI geophysical investigation parties who are applying 30 years' experience in earth sciences to measurement of earth's land and sea physical characteristics. Fundamental to modern defense, earth sciences have extended their traditional role in petroleum exploration to gathering information vital to programs in underwater warfare, missile accuracy, seismic communications, detection/surveillance of nuclear threats, and design of underground defense structures.

TI's talent for these "down-to-earth" studies dates back to 1930, when the company was formed as Geophysical Service Inc.—the first independent company to perform seismic reflection surveys for oil exploration. And from TI's early need for sensitive electronic seismic equipment

came an engineering skill that was applied to Anti-Submarine Warfare systems in World War II, and resulted in TI's Apparatus division becoming a leading ASW equipment manufacturer today. This rare blending of earth sciences knowledge and experience with skills in systems and components gives TI a unique capability in space-age technology.

The founding company (GSI) is now the geophysical exploration arm of TI's Geosciences & Instrumentation division, and these capabilities are extended to governmental agencies by the Geosciences department. Designing and manufacturing advanced instruments and systems in wide use by these parties and others is the Instrumentation Product group, with such products as WORDEN\* gravity meters, EXPLORER\* seismograph systems and seisMAC® seismic computers.

For more about geosciences at TI, write to Central Merchandising for Bulletin DM-101.

\*Trademark of Texas Instruments

**TEXAS**  **INSTRUMENTS**  
INCORPORATED

**HEADQUARTERS:** 6000 LEMMON AVENUE, DALLAS 9, TEXAS. **PLANTS:** ATTLEBORO, MASS. • DALLAS, TEXAS • HOUSTON, TEXAS • VERSAILLES, KY. • ELIZABETH SOUTH, AUSTRALIA • ALMELO, HOLLAND • AVERSA, ITALY • BEDFORD, ENGLAND • BONNEVILLE, FRANCE • BUENOS AIRES, ARGENTINA • MEXICO CITY, MEXICO

**OFFICES IN 75 PRINCIPAL CITIES OF THE WORLD**

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers



ANNOUNCING A TRANSLATION OF  
**GEODESY AND CARTOGRAPHY**

(*Geodezia i Kartografiya*)

The leading monthly journal of Geodesy and Cartography in the USSR is being translated and published in an English edition, *for the year 1959*, by the American Geophysical Union, aided by a grant from the National Science Foundation.

Subscription price, \$20.00 for the volume of 12 numbers

Send subscriptions to

**AMERICAN GEOPHYSICAL UNION**

1515 Massachusetts Avenue, N.W.  
Washington 5, D. C., U.S.A.

Other publications of the AGU include

- Journal of Geophysical Research** (monthly), \$20.00 per calendar year (approximately 400 pages anticipated for 1960)
- Transactions** (quarterly), \$4.00 per calendar year (approximately 600 pages anticipated for 1960)
- Geophysical Monograph Series** (occasional volumes), comprising papers too long for the *Journal*, or containing numerous papers of symposia and conferences related to a particular field of interest or project; Number 5 (issued in 1960) *Physics of Precipitation* (Proceedings of the Cloud Physics Conference, Woods Hole, Massachusetts, June 3-8, 1959, 435 pages, \$12.50)
- Izvestiya**, Academy of Sciences, USSR, Geophysics series, English edition (monthly), \$25.00 per calendar year, available for 1957, 1958, 1959, and on subscription for 1960

**BULLETIN (IZVESTIYA), ACADEMY OF SCIENCES, U.S.S.R.**  
**GEOPHYSICS SERIES**

Subscriptions for 1960 series now available

This monthly Russian publication, perhaps the leading journal of Geophysics of the U.S.S.R., is being translated and published in an English edition for the year 1960 by the American Geophysical Union. The twelve numbers in Russian cover about 2000 pages. Published with the aid of a grant from the National Science Foundation.

Send subscriptions now to

**AMERICAN GEOPHYSICAL UNION**

1515 Massachusetts Avenue, N.W.  
Washington 5, D. C., U.S.A.

Subscription rates: \$25.00 for the volume of 12 numbers (\$20.00 to individual members of AGU subscribing for personal use)

Numbers will be mailed as issued.

The English edition of this publication for 1957 has been translated and published for the American Geophysical Union by Pergamon Press. This volume may be ordered through the American Geophysical Union at a price of \$25.00. The 1958 and 1959 series are available at a price of \$25.00 for each volume of 12 numbers. Titles and authors of the papers contained in the series have been published in the 1959 and 1960 issues of the *Transactions*, AGU.

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertise



# NO NEW TITLES FOR GEOPHYSICISTS!

## THE EXPLORATION OF SPACE

*edited by Robert Jastrow*

space age symposium sponsored by the National Academy of Sciences, the National Aeronautics and Space Administration, the American Physical Society with contributions by 13 world-renowned scientists:

Fred L. Whipple  
Thomas Gold  
Eugene Parker  
James A. Van Allen  
N. C. Christofilos  
Homer E. Newell  
Gerard P. Kuiper  
Harold C. Urey  
G. de Vaucouleurs  
Herbert Friedman  
Leo Goldberg  
Robert Jastrow



## THE EXPLORATION OF SPACE

a thorough, analytical study of the problems, future capabilities, and recent achievements of space physics as reported by these distinguished geophysicists and astronomers at their 1959 conference.

Van Allen discusses the geomagnetically trapped corpuscular radiation; Christofilos reports on "The Argus Experiment;" de Vaucouleurs, on Mars and Venus; Goldberg, astronomy from satellites and space vehicles; Whipple, on solid particles in the solar system.

Other vital topics include rocket astronomy, plasma and magnetic fields in the solar system, the U.S. space exploration program, and the U.S.S.R. report of an active volcano on the moon.

The price is \$5.50.

## HANDBOOK OF GEOPHYSICS

*revised edition*

Prepared by the U.S.A.F.  
Geophysics Research Directorate

U.S.A.F. experts compiled this vital reference to meet the critical need for geophysical and astrophysical data in the design of aircraft, missiles, and satellites.

First published in 1957 for the exclusive use of air force prime contractors, the **Handbook** has just been revised and brought completely up to date for general distribution. This unique volume contains 680 pages (8" x 11") with 21 chapters covering the following major topics:

- Atmosphere
- Temperature
- Wind
- Precipitation
- Clouds
- Meteors
- Geomagnetism
- Terrestrial Surface Parameters
- Visibility
- The Ionosphere
- Thermal and Cosmic Radiation
- The Sun
- Jet Aircraft Condensation Trails



## HANDBOOK OF GEOPHYSICS

includes nearly 500 figures, over 200 tables, and numerous diagrams, plates, and other illustrations. Much of the information is presented in the form of probability and frequency tabulations.

Sixty scientists contributed to this detailed volume, gathering material from satellite, rocket, and solar observations; balloon and aircraft

flights; laboratory experiments; and arctic expeditions.

The price is \$15.00.

Order these important new books from your nearest bookstore or write Dept. JGR-1

*The Macmillan Company* 60 Fifth Ave., New York 11, N. Y.

In Canada: Brett-Macmillan Ltd., Galt, Ontario

Prices subject to change.

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers



# GEOPHYSICAL MONOGRAPH SERIES

## AMERICAN GEOPHYSICAL UNION

1515 MASSACHUSETTS AVENUE, N.W.

WASHINGTON 5, D. C., U.S.A.

**Antarctica in the International Geophysical Year**—Geophysical Monograph No. 1 (Publication No. 462, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 56-60071; 133 pp. and large folded map of the Antarctic, 1956, 7" x 10", \$6.00. Contains 16 papers by various American authorities on the Antarctic under the headings: General, Geographic and Meteorological, Geological and Structural, Upper Atmospheric Physics, and Flora and Fauna. Map (41" x 41") compiled by the American Geographical Society. Introduction by L. M. Gould.

**Geophysics and the IGY**—Geophysical Monograph No. 2 (Publication No. 590, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 58-60035; 210 pp., 1958, 7" x 10", \$8.00. Contains 30 papers by leading American authorities under the headings: Upper Atmospheric Physics, The Lower Atmosphere and the Earth, and The Polar Regions. Preface by Joseph Kaplan.

**Atmospheric Chemistry of Chlorine and Sulfur Compounds**—Geophysical Monograph No. 3 (Publication No. 652, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 59-60039; 129 pp., 1959, 7" x 10", \$5.50. Based on a symposium held jointly with the Robert A. Taft Sanitary Engineering Center of the U. S. Public Health Service in Cincinnati in November, 1957. Contains 23 papers (some as summaries) with discussion. Preface by James P. Lodge, Jr.

**Contemporary Geodesy**—Geophysical Monograph No. 4 (Publication No. 708, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 59-60065; 96 pp., 7" x 10", 1959, \$5.50. Based on a Conference held at Cambridge, Massachusetts, in December 1958 jointly by the AGU with the Smithsonian Astrophysical Observatory and the Harvard College Observatory. Contains 14 papers by leading authorities, with verbatim discussions on topics ranging from classical geodesy to trilateration by underwater sound to space navigation in the solar system. Edited by Charles A. Whitten and Kenneth H. Drummond.

**Physics of Precipitation**—Geophysical Monograph No. 5 (Publication No. 746, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 60-60010; 435 pp., 7" x 10", 1960, \$12.50. Based on a Conference held at Woods Hole, Massachusetts, in June 1959. Contains 48 papers by leading authorities, with verbatim discussions on topics ranging from planetary-scale phenomena to microanalysis including hail formation and precipitation control. Edited by Helmut Weickmann.

Postage is to be added to prices shown unless payment accompanies order. Quantity discounts (count each Monograph separately): 5-19 copies, 10%; 20-49 copies, 15%; 50 or more copies 20%.

### Purchase Order

### TO AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Avenue, N.W., Washington 5, D. C., U.S.A.

Please enter our order for the following:

|  |          |
|--|----------|
| _____ copies of Geophysical Monograph No. 1, at \$6.00*  | \$ _____ |
| _____ copies of Geophysical Monograph No. 2, at \$8.00*  | \$ _____ |
| _____ copies of Geophysical Monograph No. 3, at \$5.50*  | \$ _____ |
| _____ copies of Geophysical Monograph No. 4, at \$5.50*  | \$ _____ |
| _____ copies of Geophysical Monograph No. 5, at \$12.50* | \$ _____ |

- ☐ Payment of \$ \_\_\_\_\_ is enclosed.
- ☐ Please send invoice, adding postage charges.
- ☐ Enter our standing order for \_\_\_\_\_ copies of subsequent Geophysical Monographs at the special prepublication rates, e.g., prepublication rate for Monograph No. 4 for non-members was \$4.00, payment in advance, or \$4.75 (plus postage) on invoice.

\* List price is net for quantities up to four; see above for discounts on quantity purchases. Special discounts to members.

Typed name \_\_\_\_\_ Signature \_\_\_\_\_

Address \_\_\_\_\_

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertise

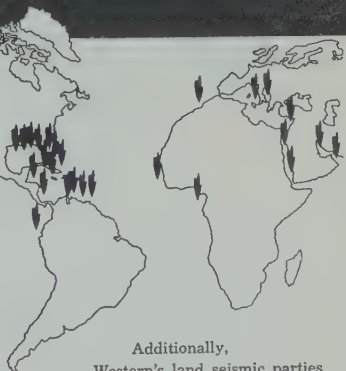




## Western Conducts More Marine Seismic Surveys Than All Other Contractors Combined!

Since 1954, Western Geophysical Company's techniques, instrumentation and professional excellence have won over half of *all* contracts for offshore seismic exploration.

The Western Marine Seismometer, a low-impedance, pressure-sensitive device, is unequalled for high output, low-frequency response, and ruggedness. Western records the entire spectrum from  $\frac{3}{4}$ -cps to 10-Kc, plus. Western's high speed cable permits normal production of up to 200 recording locations a day. The story of Western's current activities—"Panorama of a World Wide Organization"—will be sent to you upon request.



Additionally,  
Western's land seismic parties  
and gravity crews  
are at work on four continents.

*Western*

**GEOPHYSICAL COMPANY**

A SUBSIDIARY OF LITTON INDUSTRIES



Principal Office: 933 NORTH LA BREA AVENUE, LOS ANGELES 38 CALIFORNIA

AFFILIATE AND REGIONAL OFFICES THROUGHOUT THE WORLD

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers

## Advances in Space Science

Edited by FREDERICK I. ORDWAY, III

*George C. Marshall Space Flight Center,  
National Aeronautics and Space Administration, Huntsville, Alabama*

Editorial Advisory Board: WERNER VON BRAUN, FREDERICK C. DURANT, III, EUGEN SÄNGER, LESLIE R. SHEPHERD, GEORGE P. SUTTON, AND ÉTIENNE VASSY

Volume 1, 1959, 412 pp., *illus.*, \$12.00

Volume 2, July 1960, 450 pp., *illus.*, \$13.00

### CONTENTS

**Experimental Physics Using Space Vehicles**

By CHARLES P. SONETT

**Tracking Artificial Satellites and Space Vehicles**

By KARL G. HENIZE

**Materials in Space**

By FREDERICK L. BAGBY

**Plasma Propulsion Devices**

By MORTON CAMAC

**Electrostatic Propulsion Systems for Space Vehicles**

By ERNST STUHLINGER and  
ROBERT N. SEITZ

**Attitude Control of Satellites and Space Vehicles**

By ROBERT E. ROBERSON

AUTHOR INDEX—SUBJECT INDEX.

Volume 3, 1961, *in preparation*

## Advances in Geophysics

Edited by H. E. LANDSBERG *U. S. Weather Bureau*

and J. VAN MIEGHEM *Royal Belgian Meteorological Institute*

Volume 7, November 1960, about 350 pp., *illus.*, approx. \$10.00

**Developments in Controlled Experiments on Larger Scale Geophysical Problems**

By DAVE FULTZ

**Atmospheric Tides**

By MANFRED SIEBERT

**Generalized Harmonic Analysis**

By J. VAN ISACKER

**Temperature and Wind in the Lower Stratosphere**

By H. A. PANOFSKY

**Arctic Meteorology (A Ten-Year Review)**

By A. D. BELMONT

**Phase Relations of Some Rocks and Minerals at High Temperatures and High Pressures**

By GEORGE C. KENNEDY

AUTHOR INDEX—SUBJECT INDEX.

## Physics of the Upper Atmosphere

Edited by J. A. RATCIFFE *Cavendish Laboratory*

1960, 586 pp., *illus.*, \$14.50

THE AUTHORS, leaders in their fields, have written in detail about such matters as the composition of the atmosphere and its extensions into outer space, the effect upon it of the sun's radiations, and the light which it emits when it is undisturbed and when it is disturbed during an auroral display. They discuss the deductions which can be based upon observations made with the help of rockets and artificial satellites, studies of the earth's magnetic field, interpretation of radio wave reflection from the ionosphere, and indications from meteor trails and aurorae. Available results of work done during the International Geophysical Year are included in a "postscript" chapter.

## Weather Forecasting for Aeronautics

By JOSEPH J. GEORGE *Eastern Air Lines*

1960, 673 pp., 650 illustrations, \$15.00



ACADEMIC PRESS, New York and London

111 FIFTH AVENUE  
NEW YORK 3

17 OLD QUEEN STREET  
LONDON, S.W.1



# Journal of GEOPHYSICAL RESEARCH

VOLUME 65

SEPTEMBER 1960

No. 9

## SYMPOSIUM ON THE EXOSPHERE AND UPPER *F* REGION

### Summary of the Proceedings

C. O. HINES

*Defence Research Board  
Ottawa, Canada*

**Introduction.** A decade ago it was commonly believed that the hydrostatically supported atmosphere of the earth decreased in density to very low values within a few hundred kilometers of the earth's surface. This view was based on assumed thermodynamic temperatures of only a few hundred degrees, applied to the oxygen and nitrogen atoms which dominate at the heights concerned. Similar considerations indicated a corresponding confinement of the overlying exosphere—of the region, that is, where the constituent neutral atoms are free from significant collisions with one another, and where instead their motions follow ballistic trajectories. As a consequence of these conclusions, the total extent perceived for the earth's sensible gaseous mantle was limited to heights of a few thousand kilometers at most.

These views are, of course, now totally revised, and regions of the earth's atmosphere out to several earth radii are currently subject to intensive study by a wide variety of techniques. In order to provide a forum for the discussion of recent results and even newer speculations, a symposium on the exosphere and upper *F* region was sponsored by Commission III of the United States National Committee of URSI at the 1960 spring meeting in Washington. A number of formal presentations were given, and they were accompanied by several informal contributions and much discussion. The present paper constitutes a report on the proceedings.

The prepared contributions fell roughly into four broad categories, all concerned primarily with quiescent conditions and with major constituents. First, the general features of the exosphere and upper *F* region were discussed, and the various neutral particles and positive ions that dominate in successive height ranges were debated. There followed a number of papers concerning electron densities in and above the *F* region as revealed by magnetoionic refractive effects imposed on radio transmissions from rockets and artificial satellites. Refractive effects can also be applied to a study of the integrated electron content, even when detailed densities cannot be derived, and papers pursuing this line of attack came next. Finally, a group of papers treated recent developments in the study of electron densities and electron-ion temperatures as measured by means of the incoherent scatter of radio waves from individual electrons. These four groups of papers, together with the corresponding discussion and related material, are reviewed in the following four sections and summarized in a fifth.

In the review, no attempt is made at an encyclopedic abstract of the proceedings. Instead, the salient and interrelated points of the entire discussion are emphasized at the expense of detail, and are even expanded upon in certain cases. In order that some of the more detailed aspects of the proceedings should be conveniently available, however, the various contributors were

invited to provide summaries of their material, or full manuscripts if publication was imminent anyway. The papers that were forthcoming before the publication deadline are printed immediately after this report.

*General composition of the exosphere and upper  $F$  region.* In an invited introductory paper, F. S. Johnson reviewed recent developments in our understanding of the regions concerned—of their composition, of their origin, and of the physical processes that dominate them. One overriding factor contributing to our current understanding is the recognition that high temperatures prevail in and above the  $F$  region—a temperature of perhaps 1250° K is established near the top of the  $F_1$  region and maintained at greater heights (so long as 'temperature' has a meaning). It is to this fact, and to the increased emphasis now placed on the role of hydrogen at great heights, that we owe our current explanation of the extended terrestrial atmosphere. The high temperature is itself not fully understood, although recent views on the absorption of photoionizing radiation look promising, and hydromagnetic heating provides another possible source of the requisite energy.

In the model described by Johnson, interatomic collisions become negligible at a height of 550 km, which is therefore taken as the base of the exosphere. Atomic oxygen (and to some extent nitrogen) diffusing upward from the ionospheric heights below establishes an equilibrium distribution which dominates the exosphere to a height of perhaps 1200 km. Hydrogen atoms, also diffusing upward from the  $F$  region, form a minor constituent below this height; but their concentration decreases with increasing height much more slowly than that of the more massive oxygen atoms, and at 1200 km or so they come to dominate the particle distribution.

The distribution of positive ions is determined by somewhat different considerations. Below and in the neighborhood of the  $F$ -layer maximum, ionization and recombination or attachment play vital roles. Diffusion is important in the  $F$  region, and becomes dominant at levels somewhat above the  $F_2$  maximum; a diffusive equilibrium is established in the exosphere, and it is governed by hydrostatic as well as ballistic processes because interion collisions never become negligible. In these circumstances, the ions tend to have

greater scale heights (defined as  $H_i = -N_i^{-1} dN_i/dh$ , say, where  $N_i$  is the ion particle density and  $h$  is height) than the corresponding neutral particles. This is because the diffusion is ambipolar, being enforced in part by electrostatic fields set up by differential electron-ion diffusion, and the temperature that is effective in determining hydrostatic equilibrium conditions is then a sum of the electron and ion temperature,  $T_e + T_i$ . When  $T_e = T_i = T_n$  (the temperature of the neutral particles), as Johnson took it to be, the ion scale height is just twice the scale height  $H_n$  of the corresponding neutral constituent. The necessity of distinguishing clearly between  $T_e$  and  $H_n$  in any account of observational results was pointed out by Johnson and was reemphasized both by him and by other contributors in the course of the proceedings.

In the model presented by Johnson, ionized oxygen atoms provide the dominant positive-ion constituent up to a height of about 1600 km. Protons, produced in the upper  $F$  region by charge exchange between hydrogen atoms and oxygen atoms and oxygen ions, form a minor but ever-increasing fraction of the ionization up to this height, and at 1600 km come to dominate. At 2 earth radii, they become even more numerous than neutral hydrogen and so provide the dominant heavy particles of the outer regions. Their scale height can become infinite at several earth radii, or even negative, because of the centrifugal force imposed on them through a requirement for isorotation with the geomagnetic field lines.

Irregularities of ionization are observed throughout the exosphere, and are found to be columnar in shape, aligned along the geomagnetic field lines. Such irregularities could be produced directly by diffusion, as Johnson pointed out, should they be generated locally at, say, 600 km. It may be added that electrostatic and hydromagnetic forces acting from substantially lower levels could produce a similar columnar irregularity.

S. F. Singer, in the following paper, discussed many of the same questions from a somewhat different viewpoint. He placed special emphasis on the role of ballistic trajectories in the determination of the exospheric distribution of neutral particles. In similar vein, he considered the assignment of temperatures in the region to be somewhat inappropriate because of the



thermodynamic equilibrium. If temperatures assigned, however, he considered that  $T_e$  could appreciably exceed  $T_i$ , and that  $H_e$  could then exceed  $2H_i$ . In response to a question on day-to-night temperature changes recorded by satellite-drag data, he commented that changes of drag were indeed observed but were probably more properly associated with changes in the atmosphere at the base of the exosphere with temperature changes within it.

Another important view expressed by Singer was that ionized oxygen atoms become the dominant constituent of the atmosphere at a height of 100 to 1200 km, and that hydrogen atoms become dominant only at some considerably greater height. This view, based on the analysis of recent satellite data, represented a change from his earlier belief. It also represented an apparent divergence from the views expressed by Johnson and summarized above. In response to a question from Johnson, he noted that the dominance of ionized oxygen could be associated with an increased  $T_e$  if it were not otherwise explained; an increased  $T_i$ , which Johnson preferred to avoid, was not required.

On further reflection, it can be seen that the discrepancy may be in part a matter of semantics. Johnson's model indicates that hydrogen atoms become the most numerous particles at about 1200 km, whereas Singer's requirement is that the ionized oxygen shall become the most important constituent; because of the large oxygen/hydrogen mass ratio, these two conclusions are not at all incompatible. Indeed, the composition illustrated by Figures 2 and 3 of Johnson's paper can be found to yield ionized oxygen as the most dense constituent between heights of 1250 to 1800 km; neutral oxygen atoms contribute a higher density immediately below this height range, and neutral hydrogen immediately above. Clearly, it is necessary to distinguish unambiguously the sense in which a 'dominant' constituent dominates.

The proton content of the exosphere was discussed by R. E. Barrington, in a paper, by Barrington and Nishizaki, treating its effect on whistler dispersion. At frequencies comparable to their gyro frequency, protons alter this dispersion by an amount that can be measured if sufficiently pure-toned whistlers are analyzed with suitable equipment. The treatment cannot

yield a proton distribution, but only a weighted integral of that distribution; it can determine uniquely any one-parameter model of the distribution, or it can be used to test any arbitrary model. A model in which protons exist only above some transitional level, and exist there to the exclusion of any other positive ions, consistently yielded a transitional level of  $1000 \text{ km} \pm 250 \text{ km}$ . Johnson's model, if converted into this equivalent form, would yield a transitional level of 1500 km. In view of the experimental difficulties and theoretical uncertainties, the discrepancy, though appreciable, is scarcely surprising; the direct confirmation of the presence of protons is far more significant.

In reference to queries, Barrington commented that propagation along a geomagnetic field line had been assumed, and that the assumption was supported by the theory of ducting along the columnar irregularities previously described. Moreover, a possible uncertainty of  $10^\circ$  of latitude in the base of the whistler path had been reduced to perhaps  $2^\circ$  by analysis of the variation of dispersion introduced by electron gyro effects, and this uncertainty was quite tolerable.

The whistler technique can also be applied to the study of the over-all ionization distribution, as was related by R. L. Smith (Smith and Helliwell). This is achieved by examination of the dispersion particularly where electron gyro effects alter it appreciably (and produce 'nose whistlers'). Models of the ionization distribution proposed by Dungey and by Johnson had been examined on this basis, and found inappropriate unless severe assumptions relating to latitude variations and/or the exospheric temperature were made. Preliminary alternative models were proposed but not discussed in detail.

*Electron densities derived from refractive effects.* Ground-based reception of radio transmissions from rockets and satellites has become increasingly popular as a means of determining local electron densities in the upper atmosphere. However, the technique, whether it involves dispersive Doppler shifts, (Faraday) polarization changes, or differential ray directions, is subject to an inherent limitation: the effects of local electron densities at the transmitter are open to a masking by changes in the content along the transmission path. These changes can result both

from time variations of the over-all content and from the motion of the path through local irregularities or along horizontal gradients of electron content. Unless suitable corrections are made, or shown to be negligible, any unusual results obtained in application are then subject to severe reservation.

Such was the case with the work of D. J. Farmer and W. A. Robinson, and to some extent with that of C. D. Graves, in which local electron densities at satellite heights were studied. Abnormally high values were found in both instances— $4 \times 10^5$  electrons/cm<sup>3</sup> at 800-km altitude on August 2, 1958, for example, reported by Farmer and Robinson, and  $2 \times 10^4$  electrons/cm<sup>3</sup> at 18,000 km on August 23, 1959, reported by Graves. These figures exceed those normally anticipated by factors of about 4 and 10, respectively. The results were considered by their authors to be significant—despite difficulties of the type indicated above which were raised and emphasized in the discussion—and it was speculated that they might be associated with increased solar activity. Graves' results were in fact obtained towards the end of a period of intense magnetic disturbance. Smith pointed out, however, that those results were not supported by whistler measurements of August 18 and 22, 1959, which showed, if anything, a decrease of the electron content (integrated, with a weighting factor, along field lines up to heights of 3.1 to 4.4 earth radii); he felt that any straightforward interpretation of the whistler data would in fact lead to electron densities at 18,000 km down by a factor of 40 or more on the figures quoted by Graves. Regardless of one's own expectations on the resolution of this particular discrepancy, it must be recognized that the extent to which solar and magnetic activity can alter the exospheric electron content remains a problem of prime importance.

Substantial corrections were necessary in places, and were applied, in the work presented by W. W. Berning. This described an electron-density profile extending up to 1500-km height, derived from a single rocket ascent. The data were fitted to a Chapman-like profile near the *F*-layer maximum, at 262 km, and there provided an estimate of 30 km for  $H_n$ . The variation at greater heights (up to 762 km) was interpreted on a somewhat hybrid basis, which as-

sumed in effect that the ionization was in diffusive hydrostatic equilibrium with a constant  $dH_i/dh$  above 262 km, and that  $H_i = 30$  km at that level. The most appropriate value for  $dH_i/dh$  was then found to be 0.45, and a value half as great was associated with  $dH_n/dh$ .

Such a procedure, though highly dubious in some counts, receives a measure of support from comments received from J. W. Wright after the symposium. Because of their relevance, they are included among the papers of the symposium proper, after this report. Wright asserts that the presence of a neutral scale-height gradient  $dH_n/dh$ , a Chapman-like distribution of ionization will once more be obtained. It may then be anticipated that  $H_i = 2H_n$  well above the *E*-layer maximum, just as in the case where  $1/H_n$  is constant, and hence that  $dH_i/dh = 2dH_n/dh$ , as supposed by Berning. Data provided by Wright indicate that  $dH_n/dh = 0.2$  at 300–450 km (over Puerto Rico), which is then in reasonable agreement with Berning's conclusion. (An even closer agreement would be obtained if Berning's linearly increasing  $H_i$  were taken to start at the value  $H_i = 60$  km, i.e.  $2H_n$ , at the reference level of 262 km.) Although these models are in conflict with the theoretical conclusion of an isothermal upper *F* region, as described by J. W. Wright, further analysis based upon them seems to be well warranted empirically.

J. S. Nisbet described data from a set of rocket flights, indicating extensive precautions that had been taken to guard against errors of the type noted previously. His data refer to various times of day and extend in season from midwinter to early autumn. One ionization profile was obtained under conditions comparable with Berning's, namely, in the period shortly after sunrise, and indicated a very similar behavior. This represented an extreme condition in the diurnal variation, however, for all other profiles showed more gradual declines at heights above the *F*-layer maximum; the opposite extreme was obtained just after ground sunset. Nisbet suggested that this behavior indicated a departure from equilibrium conditions, brought about by the time lag of the diffusion process. Such a conclusion might remove any requirement for a height variation of temperature, and so provide an attractive alternative to the interpretation adopted by Berning and Wright.



Nisbet also established from his data that the integrated electron content between fixed levels in the upper  $F$  region varied diurnally in somewhat the same manner as the ionization density at the  $F$ -layer maximum, although some significant differences were also noted. As far as the two variations were in direct proportion, the result was consistent with a Chapman-like profile with constant scale height, as was remarked by Wright in discussion. The apparent scale heights varied appreciably, however, and discrepancies between the two variations might be a direct consequence of this.

A further point made by Nisbet was emphasized by subsequent discussion. His rocket data clearly reveal that, at night, the height of the  $F$ -layer maximum ( $h_{\max}F$ ) was considerably greater than would be deduced from a true-height analysis of ground-based ionosonde records. Wright commented that the discrepancy appeared to be in the wrong sense if it was due to the neglect of low-lying low-density ionization. In Nisbet's view, which was supported strongly by R. R. Schmerling, the discrepancy was due in part to an approximation in the analysis and in part to the occurrence of very broad ionization layers. In any event, Nisbet was able to draw the conclusion that estimates of the ratio 'total electron content above  $h_{\max}F$  to content below' are likely to be seriously in error at night through incorrect 'true-height' determinations of  $h_{\max}F$ . An alternative procedure was proposed: the ratio be determined for some fixed dividing height lying below  $h_{\max}F$  rather than at the variable and uncertain  $h_{\max}F$  itself.

The integrated electron content derived from refractive effects. The dispersive property of ionospheric propagation was employed by W. J. Ross to determine the electron content integrated from ground to satellite heights, and his results provided new information on the variation of electron content with time of day, season, and magnetic activity. The diurnal variation was consistent with that found by Nisbet: the content steadily increased from sunrise to sunset, despite the decrease of density at the layer maximum during the late afternoon hours. This result was clearly indicative of nonequilibrium conditions during most of the day and implied the necessity of a revision in the theoretical description of upper  $F$ -region profiles.

The seasonal variation showed the same decline of content from winter to summer that characterizes the density at the layer maximum, despite a tendency towards a thickening of the layer during summer months. In discussion, it was suggested by T. E. Van Zandt that the severe decline (amounting to a factor of 2 or 3, despite an increase of 2 in the effective ionizing radiation) could hardly be explained except by a seasonal variation of 4 or 5 in the linear recombination coefficient of the  $F$  region. No explanation of this variation was forthcoming.

The subsatellite electron content was found by Ross to vary inversely with the magnetic  $K_p$  index averaged over the preceding 24 hours. This conclusion tends to conflict with explanations, based on abnormal solar control, of the results obtained by Farmer and Robinson and by Graves, although it refers of course to much lower altitudes near the  $F$ -layer maximum and so is by no means decisive.

The measurement of total electron content, as revealed by the Faraday rotation of plane-polarized radio signals reflected from the moon, was reviewed by R. B. Dyce and summarized in terms of the ratio 'content above  $h_{\max}F$ /content below.' This ratio had been found by previous workers to decrease from 4 or 5 at night to perhaps 3 during the day, while more recent measurements made by R. A. Hill and himself had indicated a reverse trend from 1.5 at night to 2.5 during the day. G. H. Millman, in a following paper (Millman, Sanders, and Mather), reported other recent determinations in which the night values again exceeded those of the day, being perhaps 6-8 and 3-5, respectively. The same data reconfirmed the conclusions of Nisbet and Ross that the electron content continues to increase until the hours of sunset.

Several important points were established or re-emphasized by these papers and by the discussion accompanying them. First, it was clear that the ambiguities of the single-frequency Faraday technique rendered the significance of its conclusions somewhat uncertain; the resolution of the ambiguities by means of assumptions about the medium under study is at best unsatisfactory, and two-frequency observations are to be encouraged whenever possible. At the very least, information on the sense of polarization rotation should be retained.

Again, uncertainties about the content below the  $F$ -layer maximum provide a potential source of error in describing the results. These uncertainties arise from at least two causes, both of which have already been indicated: at night, at least, the height of the  $F$ -layer maximum is not accurately defined by 'true-height' analyses, and at all times horizontal gradients of ionization may render the data of the nearest 'true-height' station inapplicable to the measurements of the lunar radar. This second point was particularly pertinent to the data of Millman and co-workers, since horizontal gradients are severe at the latitudes concerned (between Trinidad and Puerto Rico).

Finally, it was emphasized by Millman that the common practice of removing the magnetic-field factor from the integrand, when integrating the theoretical Faraday rotation, introduces a source of error. The error is not great for propagation at small zenith angles but can become important when the moon lies close to the horizon.

In later discussion, C. G. Little outlined a proposed application of the multi-frequency phase and Faraday-rotation techniques in an experiment employing a '24-hour satellite.' Since such a satellite would remain approximately stationary in the observer's sky, it should be possible to obtain continuous and extremely sensitive measurements of integrated electron content, free from ambiguities and sources of serious error. Details of the proposal were published by Garriott and Little in the July issue of this journal.

*Electron-ion densities and temperatures derived from incoherent backscatter.* The concluding formal presentations gave strong evidence of the very valuable role that can now be played by measurement of incoherent backscatter from the ionosphere, a technique first proposed by W. E. Gordon and implemented by K. L. Bowles. The method, employing a high-power radar operating at a frequency above the  $F$ -layer critical, measures the integrated scatter from individual electrons.

T. E. Van Zandt (Van Zandt and Bowles) described his recent analysis of data obtained by Bowles at 41 Mc/s. The upward decline of ionization densities above the  $F$ -layer maximum was examined at times of differing solar activity,

and found to be less rapid when the activity was high (apparently in conflict with the conclusion of Ross). Temperature variations over the range  $1000^\circ$  to  $1500^\circ\text{K}$  could be derived from the observed scale heights. As the analysis was subject to a variety of uncertainties that had not yet been fully resolved, the conclusion was presented as a preliminary one only.

Extremely valuable results were presented by V. C. Pineo, (Pineo, Kraft, and Briscoe) obtained by application of the backscatter technique at 440 Mc/s. These showed striking improvements over data reported earlier, both in the distribution of ionization with height and in the spectral broadening introduced by thermal motions of the charged particles. The decline of ionization in the upper  $F$  region proceeded with a scale height,  $H_i$ , of about 120 km in one example presented, and this scale height appeared to increase slightly as the exosphere was entered. The frequency spectrum of signals returned from the  $F$  region at 330 km was clearly flat-topped, rather than Gaussian in shape, and it exhibited a width slightly more than twice that of the  $E$ -region return from 120 km. The square of ratio—about 5—might then be expected for the ratio of the ion temperatures at the two heights, if equal ion masses were assumed. The scale height for atomic oxygen,  $H_n = H_i/2 = 60$  km in the upper region, would correspond to a temperature

$$T_n \simeq 1100^\circ\text{K}$$

An absolute interpretation of the spectral broadening in terms of temperature was rendered possible by formulas presented by J. Fejer, which had also been derived independently by J. P. Dougherty and D. T. Farley and E. Salpeter. The formulas determined the scattering to be expected from an ionized gas in thermal equilibrium, in a certain limiting case which was applicable, and predicted the flat-topped spectrum obtained by Pineo and co-workers. Indeed, it predicted two slightly raised shoulders on the spectral profile which were observed, perhaps because of the small amount of bandwidth smearing known to be present. The best fit of Fejer's formula to the observed spectrum

<sup>1</sup> Note added in proof: Better data, published in this issue by Pineo, Kraft, and Briscoe, exhibit the shoulders.



m yielded an apparent ion temperature

$$T_i \simeq 1400^\circ\text{K}$$

substantial agreement with the values quoted above for  $T_e$ . The effect of a significant increase in  $T_e$  over  $T_i$  would be to enhance the moulders of the spectrum appreciably (Fejer, private communication), and hence no support to Singer's suggestions of high  $T_e$ 's could be drawn from the observations now available.

The congratulations of the audience went to those who had pioneered in the development of this most promising technique, together with encouragement for its further application.

*Summary.* Many points of summary have already been made in the preceding sections. Little can be added at this point that was not said by Johnson in an informal summary at the conclusion of the symposium itself. His remarks therefore follow:

Data on the electron distribution above the  $F_2$  maximum obtained by a variety of observing techniques have been described. Although all the techniques are clearly capable of providing useful information, it also seems apparent that uncertainties in the interpretation of the observations in some cases are giving rise to variations which are not representative of the electron distribution above the  $F_2$  peak. It is clear that effort should be expended on means of reducing the errors that creep into the analysis of the various measurements.

From a theoretical point of view, one cannot accept the concept of a temperature gradient in an exosphere. In terms of the ion distribution above the  $F_2$  peak, this means that, more than 100 or 200 km above the peak, the scale height should be constant except in the region where the transition from oxygen to hydrogen ions occurs. In this regard, the fitting of the Chapman formula to the distribution from the  $F_2$  peak to the bottom of the

transition region in the manner recommended by Wright in the January issue of this journal is entirely satisfactory since it involves the concept of constant scale height; however, the Chapman expression should not be used in and above the transition region where protons become the predominant ion. The chief value of the Chapman expression in its application to the  $F_2$  region is that it describes the electron distribution in the region where the mechanism controlling the electron concentration changes from recombination processes at the  $F_2$  peak to diffusive processes at higher levels; there is some theoretical basis for its use, based on diffusion calculations by several workers.

There is a particular need for electron concentration data in the 1000- to 1500-km altitude region at varying latitudes, as they would give a clearer picture of the proton distribution in the protonosphere or outer ionosphere. At present, a distribution can be calculated on the basis of our physical understanding of the support mechanism; however, whistler observations at high latitudes appear not to agree with the calculated distribution. Since the whistler observations do not specifically determine the electron concentrations at altitudes near the oxygen-hydrogen transition region, independent observations are especially needed there.

#### APPENDIX

The Symposium on the Exosphere and Upper  $F$  Region was held on the evening of May 4, 1960, at the National Bureau of Standards, Washington, D. C.; the chairman was C. O. Hines of the Defence Research Board of Canada.

*Acknowledgment.* While responsibility for errors in this report must rest solely with its author, the report itself could not have been written had it not been for the cooperation of the various speakers in providing summaries or manuscripts of their contributions. This cooperation is gratefully acknowledged.





# The Exosphere and Upper $F$ Region<sup>1</sup>

FRANCIS S. JOHNSON

*Lockheed Missiles and Space Division  
Palo Alto, California*

**Abstract.** The view is presented that the outer portion of the earth's ionosphere, through which radio whistlers propagate, has its origin in the earth's atmosphere rather than in interplanetary space. The outer portion of the earth's neutral-particle exosphere is dominated by neutral atomic hydrogen which is escaping steadily from the earth's atmosphere. Some of this neutral atomic hydrogen reacts with atomic oxygen ions near the base of the exosphere, giving rise to hydrogen ions. Diffusive equilibrium is the dominant factor controlling the relative concentration of ions and neutral particles near the base of the exosphere. At higher levels, the geomagnetic field exerts an influence in that it causes the ions and electrons to rotate with the earth, and the centrifugal force modifies the force of gravitational attraction appreciably. The ion electron plasma through which the radio whistlers propagate is hydrostatically supported within the magnetic tubes of force.

Before the correct interpretation of radio whistlers was put forth by Storey [1953], it seemed impossible to conceive of any significant extension to the earth's atmosphere or ionosphere to altitudes greater than 1000 or 2000 km. The interpretation of whistlers as due to propagation of signals from radio atmospherics through a medium consisting of several hundred electrons per cubic centimeter at distances of several earth radii from the earth was a very new idea, but there was other evidence for such concentrations of electrons in interplanetary space. About this time also it was recognized that there is an outstreaming of charged particles from the sun [Biermann, 1953], now known as the solar wind [Parker, 1958], which was estimated to consist of several hundred electrons per cubic centimeter near the earth's orbit. Zodiacal-light measurements made at about the same time indicated the presence of several hundred electrons per cubic centimeter in space in the vicinity of the earth's orbit [Behr and Siedentopf, 1953]. Consequently, it was natural to assume that the ionized medium through which the radio whistlers propagate was, in some way, of solar origin.

This viewpoint was probably presented in its most complete form by Chapman [1957], who viewed the solar corona as extending beyond the earth's orbit with a concentration of several hundred protons and electrons per cubic centimeter. The effect of the earth's magnetic field in excluding charged solar particles from the vicinity of the earth was recognized, but it was argued that the solar plasma would gradually penetrate the magnetic field and, hence, finally fill the space through which the radio whistlers propagate with a plasma of solar origin. In this view the outer ionosphere was clearly dominated by the extended solar atmosphere.

It is now rather clear that the gas within the earth's magnetic field is of telluric origin and that it is simply the extreme extension of the earth's atmosphere and ionosphere resulting from the presence of atomic hydrogen in the ionosphere [Johnson, 1959a]. The first definite experimental indication of atomic hydrogen in the upper atmosphere was the nighttime observation in rockets above the  $E$  region of hydrogen-Lyman- $\alpha$  radiation [Kupperian, Byram, Chubb, and Friedman, 1958]. These data can be interpreted as indicating the presence of a vast cloud of atomic hydrogen around the earth; we will call this hydrogen cloud the telluric hydrogen corona. Figure 1 shows the distribution of atomic hydrogen in the vicinity of the earth deduced on the basis of the nighttime Lyman- $\alpha$  observa-

<sup>1</sup> This paper was presented at the Symposium on the Exosphere and Upper  $F$  Region sponsored by Commission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.

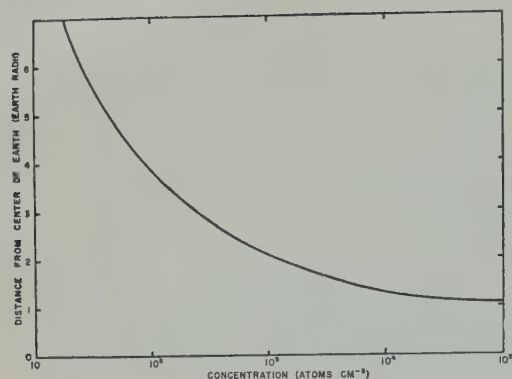


Fig. 1. The distribution of atomic hydrogen near the earth.

tions [Johnson, 1959b.] This curve was first presented at the 1959 spring meeting of the American Geophysical Union in Washington. The presence of such a telluric hydrogen corona has since been confirmed by means of highly resolved spectra of the solar Lyman- $\alpha$  line obtained at the Naval Research Laboratory [Purcell and Tousey, 1960]; these spectra indicate the presence of slightly less hydrogen than the amount shown in Figure 1. Also a slight modification of the shape of the distribution shown in Figure 1 is indicated by detailed calculations made by Öpik and Singer [1959] and by Johnson and Fish [1960].

Figure 2 shows the distribution of atomic hydrogen in the ionosphere [Johnson, 1959b] on the basis of the analysis of the nighttime Lyman- $\alpha$  observations. The distribution above 120 km is a diffusive equilibrium distribution; there is evidence that the distribution of other atmospheric constituents is under the control of diffusive processes at altitudes as low as 105 km [Meadows and Townsend, 1960]. Atomic hydrogen diffuses up through the atmosphere and escapes from the exosphere at such a rapid rate, however, that its distribution departs from diffusive equilibrium below about 120 km. The base of the exosphere is at an altitude of about 550 km, and, even at this altitude, hydrogen is a minor constituent of the atmosphere. Because of the very slow decrease of hydrogen concentration with altitude, atomic hydrogen becomes the predominant atmospheric constituent above about 1200 km; above this altitude, the atmospheric density falls very slowly with height. The

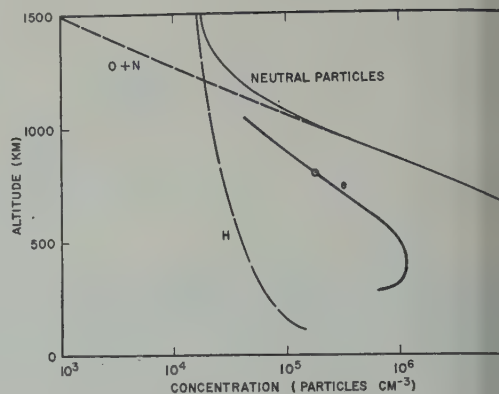


Fig. 2. The distribution of atomic hydrogen and heavier atmospheric constituents in the ionosphere. The electron distribution is also shown.

temperature in the exosphere is about 1250° although, far from the earth, the velocity distribution of the escaping hydrogen becomes non-Maxwellian and the temperature is not well defined [Öpik and Singer, 1959]. Even here, however, the average energy of the hydrogen atoms is, within about 30 per cent, the same as in the lower exosphere.

An accidental coincidence of ionization potential for hydrogen and oxygen leads to the probability that charge exchange may occur between hydrogen atoms and oxygen ions in the ionosphere, as suggested by Dungey [1955]. Recently, R. F. Stebbings (private communication) has measured the cross section for this charge exchange and has found it larger than the kinetic cross section. As a result of the charge exchange equilibrium that develops, it is to be expected that, above the  $F_2$  maximum, there will be increasing concentrations of ionized hydrogen resulting from charge exchange between hydrogen atoms of the telluric hydrogen corona and atomic oxygen ions in the upper portion of the  $F_2$  region. Figure 3 indicates such probable concentrations in the upper part of the ionosphere [Johnson, 1960].

The distribution of ions at levels well above the  $F_2$  maximum is governed entirely by diffusive processes, since diffusion exerts a more important influence than ionization or recombination in controlling the distribution. In a diffusive equilibrium, the ions are distributed as if they had a temperature equal to the temperature of their temperature and the electron temperature.



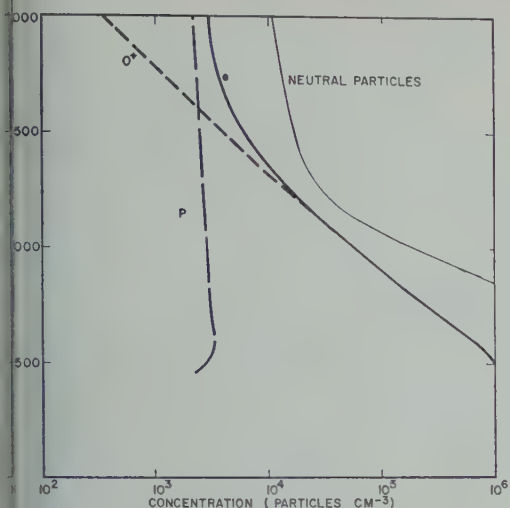


Fig. 3. The proton and  $O^+$  concentrations in the upper part of the ionosphere. The neutral-particle and electron concentrations are also shown.

ture; if these two temperatures are equal to the neutral-particle temperature above the  $F_2$  region, the scale height for ions is twice as great as for neutral particles of the same mass. The ion distributions above 550 km shown in Figure 3 have scale heights for the ion constituents which are just twice the scale heights for the corresponding neutral constituents. This is roughly in accordance with the various observations that have given an indication of the electron distribution above the  $F_2$  maximum. If the electron temperature above the  $F_2$  maximum is very high, the electron and ion concentrations would fall very slowly above the  $F_2$  region; Johnson and Johnson, however, who have studied the energy budget of electrons, conclude (in an article now in preparation) that the electron temperature above the  $F_2$  maximum must be very close to the neutral-particle temperature. It has been pointed out that the electron distribution in and above the  $F_2$  region can be reasonably fitted by the distribution computed from the analytical expression for a Chapman layer, even though the  $F_2$  region is not a Chapman layer [Wright, 1960]. The scale height which appears in the Chapman formula is that of the neutral particles, whereas the scale height of the ions and electrons well above the peak of the layer, as computed from the Chapman formula, is twice as great as the scale height of

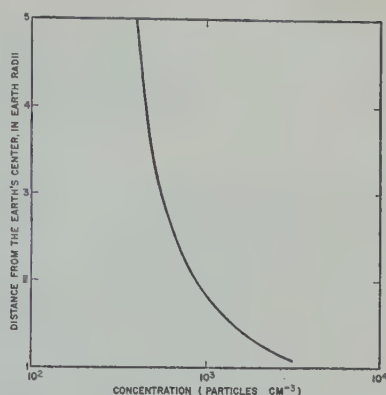


Fig. 4. The proton distribution in the protonosphere.

the neutral particles. It is important in discussing the distribution of ions above the  $F_2$  maximum that it be specified whether the quoted scale height is the ion scale height or the neutral-particle scale height in the Chapman expression which fits the observed ion distribution.

The distribution of ionized hydrogen at much higher levels is shown in Figure 4 [Johnson, 1960]. This portion of the ionosphere, consisting of protons and electrons, may be distinguished from the normal ionosphere, in which the ions consist mainly of atomic oxygen and heavier particles, by calling it the protonosphere; the designations magnetosphere and outer ionosphere have also been suggested. Throughout the altitude region shown, the ions are hydrostatically supported. The magnetic field affects the distribution of ions, but not the distribution of neutral particles, in that it links the ion-electron plasma to the lower ionosphere and requires that it partake fully of the earth's rotation, whereas the neutral particles high in the exosphere follow ballistic trajectories and maintain their angular momentum but not their angular velocity. The resulting centrifugal force on the ions modifies the gravitational attraction which enters into the hydrostatic relationship. The medium described by Figure 4 is not an ion exosphere; collisions are important throughout the altitude range shown, and the mean free path is nowhere as large as the ion scale height. It must be remarked that the distribution shown in Figure 4 applies only along magnetic field lines; at low latitudes, the curve shown applies only to the highest altitude reached by the mag-

netic field line, and for auroral and higher latitudes it does not apply at all. The distribution shown in Figure 4 has been computed from theory, and up to the present there have been too few experimental data to confirm or contradict it.

The irregularities in ion concentration which occur in the ionosphere are not clearly understood. There has been a tendency to interpret them in terms of simple turbulence, but this is difficult to understand in such a stable medium. Progress in the direction of understanding the irregularities in the  $E$  region has resulted from studies of gravity waves by *Hines* [1959], and there is reason to hope that this problem will soon be solved. The irregularities above the ionosphere, as deduced from whistler observations, seem to be columnar in nature, aligned along the magnetic field [*Smith, Helliwell, and Yabroff*, 1960]; the collision frequencies in the protonosphere are such that this is to be expected, provided that an irregularity can be produced in the first place. That is, a 10 per cent anomaly in ion and electron concentration near an altitude of 600 km in the upper part of the  $F_2$  region should be associated with a similar anomaly all along the magnetic field line passing through the postulated anomaly, even to the extent of producing a similar anomaly at the top of the  $F$  region in the opposite hemisphere. The mechanism capable of producing such an anomaly is not at present understood.

The ion and electron concentrations in the vicinity of the  $F_2$  maximum and below are under the control of recombination processes, diffusion being too slow in this altitude region to affect the distribution of ions. The ion species actually present in this region result from charge-exchange and ion-atom-exchange reactions which can occur among the various atmospheric constituents. In the  $F_2$  region and above,  $O^+$  has been observed to be the predominant ion, and a few  $N^+$  ions have also been observed, particularly at the higher levels.

The atmospheric temperature in the exosphere was required in calculating the distributions shown in Figures 2 and 4 and was determined from atmospheric density data obtained from the decay of satellite orbits. The density determinations in the upper atmosphere can be interpreted only in terms of a marked rise in

temperature with altitude in the  $E$  region. A best fit with the observed data is obtained with a temperature distribution which rises very rapidly in and immediately above the  $E$  region and becomes isothermal with a value near 1250°K above the  $F_1$  region. Such a distribution is indicative of a heat source in the  $E$  region. In the past, it has frequently been assumed that the heat source which keeps the  $E$  region warm arises through absorption of solar ultraviolet radiation [*Johnson*, 1958]. It is not clear that this particular heat source is adequate, however; some other source may be required to explain the warmth of the ionosphere. A promising though tentative source is the absorption of hydromagnetic wave energy [*Decker*, 1959]. This source would deposit energy in the  $F_1$  region as required, but necessitates the existence above the  $F$  region of hydromagnetic waves with amplitudes of about 100  $\gamma$ , roughly in the 1-cps frequency region. Observation of such amplitudes would leave little doubt that this is an important heat source in the ionosphere.

The existence of Van Allen radiation in the region through which radio whistlers propagate tempts one to associate the trapped particles with whistler propagation, but no relation has been established. The Van Allen particles make up a very minor part of the particles present in this region, and there seems to be no interaction between the trapped high-energy particles and the hydrostatically supported particles.

It now appears clear that the outer ionosphere or protonosphere is a natural extension of the normal ionosphere resulting from the presence of atomic hydrogen in the ionosphere. The ionizing radiation that produces it is the same as that which produces the  $F$  region. The protons which fill the earth's magnetic field and make up the protonosphere are formed near the base of the neutral-particle exosphere, where they are in chemical equilibrium with other atmospheric constituents; the proton concentrations above the base of the exosphere are under control of diffusive equilibrium and hence are distributed in accordance with a hydrostatic relationship.

#### REFERENCES

- Behr, A., and H. Siedentopf, Untersuchungen über Zodiakallicht und Gegenschein nach lichtelect-



- phen Messungen auf dem Jungfraujoch, *Z. Astrophys.*, **32**, 19-50, 1953.
- Armstrong, L., Physical processes in comet tails and their relation to solar activity, *Mém. soc. roy. sci. Liège, Collection in 4<sup>e</sup>*, **13**, 291-302, 1953.
- Capman, S., Notes on the solar corona and the terrestrial ionosphere, *Smithsonian Contribs. to Astrophys.*, **2**, 1-12, 1957.
- Cesler, A. J., Ionospheric heating by hydromagnetic waves, *J. Geophys. Research*, **64**, 397-401, 1959.
- Chapman, J. W., Electrodynamics of the outer atmosphere, *The Physics of the Ionosphere, Report of the 1954 Cambridge Conference*, The Physical Society, London, pp. 229-236, 1955.
- Cox, C. O., An interpretation of certain ionospheric motions in terms of atmospheric waves, *J. Geophys. Research*, **64**, 2210-2211, 1959.
- Donson, F. S., Temperatures in the high atmosphere, *Ann. géophys.*, **14**, 94-108, 1958.
- Donson, F. S., Telluric origin of the whistler medium, *Nature*, **184**, 1787-1788, 1959a.
- Donson, F. S., The structure of the outer atmosphere including the ion distribution above the  $F_2$  maximum, *Tech. Rept. LMSD 49719*, Lockheed Aircraft Corporation, Sunnyvale, California, April 28, 1959b.
- Donson, F. S., 'The ion distribution above the  $F_2$  maximum, *J. Geophys. Research*, **65**, 577-584, 1960.
- Johnson, F. S., and R. A. Fish, The telluric hydrogen corona, *Astrophys. J.*, **131**, 502, 1960.
- Kupperian, J. E., Jr., E. T. Byram, T. A. Chubb, and H. Friedman, Extreme ultraviolet radiation in the night sky, *Ann. géophys.*, **14**, 329-333, 1958.
- Meadows, E. B., and J. W. Townsend, Jr., IGY rocket measurements of arctic atmospheric composition above 100 km, *First Intern. Space Sci. Symposium*, North Holland, Amsterdam, 1960.
- Öpik, E. J., and S. F. Singer, Distribution of density in a planetary atmosphere, *Phys Fluids*, **2**, 653-655, 1959.
- Parker, E. N., Interaction of the solar wind with the geomagnetic field, *Phys. Fluids*, **1**, 171-187, 1958.
- Purcell, J. D., and R. Tousey, The profile of solar hydrogen-Lyman- $\alpha$ , *J. Geophys. Research*, **65**, 370-372, 1960.
- Smith, R. L., R. A. Helliwell, and I. W. Yabroff, A theory of trapping of whistlers in field-aligned columns of enhanced ionization, *J. Geophys. Research*, **65**, 815-823, 1960.
- Storey, L. R. O., An investigation of whistling atmospherics, *Phil. Trans. Roy. Soc. London, A*, **246**, 113-141, 1953.
- Wright, J. W., A model of the  $F$  region above  $h_{max}F_2$ , *J. Geophys. Research*, **65**, 185-191, 1960.

(Manuscript received May 20, 1960.)





# Structure of the Earth's Exosphere<sup>1</sup>

S. F. SINGER

*Physics Department, University of Maryland  
College Park, Maryland*

**Abstract.** The main components of the earth's exosphere are neutral oxygen, neutral hydrogen atoms, ionized oxygen, and ionized hydrogen. The position of the base of the exosphere is established at an altitude of 530 km, from an analysis of density data obtained from satellite drag observations. The relative distribution of both neutral oxygen and neutral hydrogen is derived from a theory of the exosphere. The normalization of oxygen is accomplished by means of satellite drag data. The normalization of hydrogen is accomplished by means of the Lyman- $\alpha$  technique of Johnson, Friedman, and Tousey. The distribution of ionized oxygen can be established jointly from the rocket measurements of Berning and from an analysis of radiation belt data obtained from satellites. They give a consistent picture of ionized oxygen becoming the most important constituent of the atmosphere at about 1500 km. The distribution of ionized hydrogen is only imperfectly known.

**Introduction.** In any planetary atmosphere a level of altitude is reached at which the density of the atmosphere is so low that the mean free path becomes of the order of the scale height. To a good approximation it can be assumed that above this level the mean free path in fact becomes infinite so that molecules traveling in this region, named the exosphere, would experience no collisions with other molecules. It can be easily seen, then, that molecules issuing from the base of the exosphere will describe ballistic orbits, i.e. portions of ellipses which take them out to high altitudes before they reenter the atmosphere at a different point. Most of the molecules will, in fact, have velocities high enough, greater than the escape velocity, to take them beyond the gravitational field of the planet.

For the planet earth, we find the critical level to be at  $h_c = 530$  km. That is an average value, obtained by an analysis of satellite drag data, free of the diurnal variation; it has been discussed elsewhere [Singer, 1960]. At  $h_c$ , we must have a concentration  $n_c$  so that  $n_c = 1/(4H_0a)$ ; with  $a$ , the mean atomic radius of  $6 \times 10^{-8}$  cm and  $H_0 \sim 10^7$  cm we find

$n_c \sim 4 \times 10^7 \text{ cm}^{-3}$ . This concentration mainly refers to atomic oxygen. Below about 350 km,  $N_2$  becomes an important constituent; below 220 to 230 km it predominates.

We may therefore take the 530-km level as the base of the exosphere. For more accurate considerations, one must take into account that at that level only the mean free path in the horizontal direction is equal to the scale height. In reality, the mean free path is not isotropic and is, therefore, a function of the angle which the velocity vector of the particle makes with the local vertical.

2. *Neutral oxygen.* Using the value  $n_0(O) = 4 \times 10^7 \text{ cm}^{-3}$  at the base of the exosphere, we obtain the mass concentration  $\rho(O)$  as 16  $n(O)$  atomic masses. It is this mass concentration<sup>2</sup> that is plotted in Figure 1.

According to the derivation by Öpik and Singer, the distribution of a neutral exospheric component with altitude is given by

$$\frac{\rho(y)}{y^2} = \int_0^{\pi/2} \frac{\sin \theta \cos \theta \exp(-Ea^2) d\theta}{(1 - y^2 \sin^2 \theta)^{1/2}} \cdot [(1 + 2Ea^2)\phi - b\phi'] + \frac{1}{2} \int_0^{\pi/2} \frac{\sin \theta \cos \theta \exp(-Ea^2) d\theta}{(1 - y^2 \sin^2 \theta)^{1/2}} \cdot [(1 + 2Ea^2)(1 - \phi) + b\phi'] \quad (1)$$

<sup>2</sup> When multiplying it by the unit of atomic mass we obtain the density in grams per cubic centimeter.

This paper was presented at the Symposium on the Exosphere and Upper F Region sponsored by Commission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.

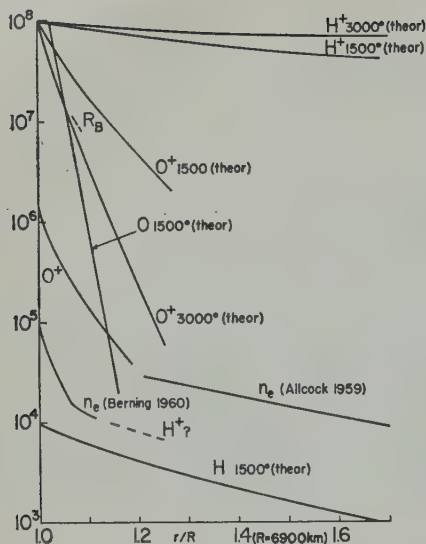


Fig. 1. Distribution with altitude of various components of the earth's exosphere. The theoretical distributions of O and H are normalized from experimental data; the ordinate is mass concentration in units of (atomic masses/cm<sup>3</sup>). Relative theoretical distributions are shown for O<sup>+</sup> and H<sup>+</sup> at two temperatures. The absolute electron-density distribution is shown as obtained by Allcock and by Berning (units are electrons/cm<sup>3</sup>). From the latter data and data of radiation-belt counting rate by Yoshida, Ludwig and Van Allen [1960] we derive the mass concentration of O<sup>+</sup> shown. The normalization for the H<sup>+</sup> distribution may be surmised from Berning's data and is indicated provisionally. (Note: The labels O<sup>+</sup>1500° and O<sup>+</sup>3000° should be interchanged. The radiation belt analysis (RB) agrees almost exactly with the curve for O but shows a smaller slope beyond 1000 km (as indicated).)

where  $y = R/r$ ,  $R$  is the reference radius =  $R_E + 530$  km = 6900 km,  $E = GMm/(RkT)$ ;  $\theta$  is the zenith angle measured from the vertical;  $a^2 = (1 - y)/(1 - y^2 \sin^2 \theta)$ ;

$$\phi = \phi(b) = 2(\pi)^{-1/2} \int_0^b e^{-x^2} dx$$

where  $b^2 = E(1 - a^2)$ ; and

$$\phi' = \phi'(b) = 2(\pi)^{-1/2} e^{-b^2}$$

The first part of equation 1 refers to the ballistic re-entry component; the second part refers to the escaping component (and is counted one-half since it contributes only on the way out).

To a good approximation equation 1 can be represented for  $30 \leq E \leq 90$  (corresponding to oxygen) by

$$n(r) = 0.98n_0(r/R)^{-0.96E}$$

For  $(r/R) \lesssim 1.10$  the accuracy is within per cent.

Therefore, using the theoretical distribution (1) and the normalization from satellite data [King-Hele, 1959] we obtain the curve O shown in Figure 1.

3. *Distribution of neutral hydrogen.* The relative concentration of neutral hydrogen with altitude has been given by Öpik and Štěrba [1959, 1960a]. The ballistic re-entry and escape component are described by the same equation 1 as used for neutral oxygen, but with the appropriate atomic mass. It should be noted that the temperature must be the same since at the base of the exosphere neutral O and H are both in thermodynamic equilibrium. Hence the appropriate value of

$$E = 4.65 \quad \text{for neutral hydrogen}$$

Empirically, we find that the hydrogen distribution cannot be accurately approximated with a single power law. For  $E = 4.65$  we have a good fit within 5 per cent in the range  $1 < r/R < 10$ :

$$n(r) = n_0[0.75(r/R)^{-5.60} + 0.25(r/R)^{-2.64}]$$

Using equation 4,

$$\int_R^\infty n(r) dr \quad \text{is found to be} \quad 0.315Rn_0$$

Here  $R = (6370 + 530)$  km. We can now compare the experimental finding of Purcell and Tuck [1960], who obtained the optical depth  $\tau$  from a study of the profile of the solar Lyman line; we therefore obtain for the thickness of neutral hydrogen  $2 \times 10^{12}$  atoms/cm<sup>2</sup> column. Hence

$$n_0(\text{H}) = 10^4 \quad \text{atoms/cm}^3$$

The absolute distribution of neutral hydrogen is plotted in Figure 1.

Some question still exists about the density to which bound ('satellite') orbits are filled; we consider that their contribution near the earth is not too important, and we will neglect them. For a more detailed discussion of



discrepancy, the reader is referred to *Brandt and Chamberlain* [1960] and *Öpik and Singer*, [1960].

When comparing our theoretical distribution, given by equation 1, with the theory of *Johnson and Fish* [1960], we find a wide difference. Although the general approach is similar, we consider the derivation of *Öpik and Singer* [1959, 1960a] to be more nearly correct. The discrepancy with Johnson and Fish affects not only the relative but also the absolute concentrations of hydrogen.

**Ionized oxygen.**  $O^+$  ions in the lower exosphere are distributed according to the barometric formula; collisions between them are not so frequent because of the large Coulomb cross sections. Hence

$$n_0 = \exp \left[ -\frac{1}{2}(GM/kT)(1/R - 1/r) \right] \\ = \exp \left[ -\frac{1}{2}E(1 - y) \right] \quad (6)$$

The factor  $1/2$  enters in, since the presence of  $e^-$  electrons effectively cancels one-half of the gravitational force by means of electrostatic repulsion. If the electron temperature is higher than the ion temperature, the factor will be  $(T_{\text{ion}} + T_e)$  in place of  $1/2$ .

Our theoretical distributions are plotted on Figure 1, for  $T = 1500^\circ\text{K}$  and for  $T = 3000^\circ\text{K}$ . In order to carry out a normalization, we must first recourse to experimental data. We make use of the directly measured electron densities given by *Berning* [1960] which were obtained from a sounding rocket. These data are shown in Figure 1. Also on Figure 1 we have plotted the electron distribution derived by *Allcock and Smith* [1960] from an analysis of the dispersion of the ordinary whistlers.

When we identify the data of *Berning* [1960] at altitudes with ionized oxygen, we can then determine the mass concentration as shown in Figure 1. Making a theoretical extrapolation, we find that oxygen becomes the most important constituent at an altitude of  $1.14R$  expressed in terms of  $R = 6370$  km, i.e. an altitude of  $\sim 1500$  km above sea level.

Similar information is obtained from the analysis of radiation-belt data [*Singer*, 1960]. The intensity of trapped protons depends on the mass of the particles, which in turn is inversely proportional to atmospheric density. The

relationship is not completely straightforward since such factors as trapping efficiency and injection coefficient have to be considered. These are obtained from a theory of the inner radiation belt.

It is quite simple then to derive the radiation-belt flux if the atmospheric density distribution is known. The analysis of the inverse problem is complicated by the fact that the earth's magnetic field is eccentric. Hence the trapped particles do not remain at a constant geocentric distance as they move around the earth, but follow a magnetic surface equidistant from the dipole.

When this analysis is carried out, the following results are obtained [*Singer*, 1960]. The radiation-belt data below 1000 km are in close agreement with and support the theoretically calculated density of atomic oxygen. At about 1000 km, the radiation-belt intensity begins to increase less rapidly, which indicates a less rapid fall-off of atmospheric density with altitude. However, because of the eccentricity of the dipole (by about  $\Delta = 436$  km from the center of the earth), the radiation-belt data reflect the state of the atmosphere over an altitude range of  $2\Delta$ , but with the higher altitudes only slightly weighted. The deviation of radiation-belt data starting at  $\sim 1000$  km indicates that another constituent with 'scale height' larger than that of O becomes more important at an altitude of 1300 to 1600 km.

Above 1000 km the radiation-belt intensity seems to assume a constant logarithmic slope which definitely does not correspond to O and agrees reasonably closely with  $O^+$  at a temperature of  $1500^\circ$ . The data are not precise enough at present to allow a more refined analysis.

**5. Ionized hydrogen.** The density distribution of ionized hydrogen near the earth is again given quite closely by a barometric formula. The variation is shown in Figure 1. The normalization here is difficult. If we adopt the point of view that the change in slope at  $\sim 1000$  km in *Berning's* data is real (as plotted in Fig. 1), it would seem to indicate a distribution of hydrogen approximately as shown there. This distribution is in fair agreement with the whistler data. It is evident, however, that more precise experimental data are required before the distribution of ionized hydrogen can be definitely established.

*Acknowledgment.* This research was supported in part by the Air Force Cambridge Research Center.

## REFERENCES

- Allcock, G. McK., The electron density distribution in the outer ionosphere derived from whistler data, *J. Atmospheric and Terrest. Phys.*, **14**, 185-199, 1959.
- Berning, W. W., A sounding rocket measurement of electron densities to 1500 kilometers, *J. Geophys. Research*, **65**, 2589-2594, 1960.
- Brandt, J. C., and J. W. Chamberlain, Density of neutral gas in a planetary exosphere, *Phys. Fluids*, **3**, 485-486, 1960.
- Johnson, F. S., and R. A. Fish, The telluric hydrogen corona, *Astrophys. J.*, **131**, 502-515, 1960.
- King-Hele, D. G., Density of the upper atmosphere from analysis of satellite orbits, *Nature*, **184**, 1267-1270, 1959.
- Öpik, E. J., and S. F. Singer, Distribution of density in a planetary exosphere, *Phys. Fluids*, **2**, 653-655, 1959.
- Öpik, E. J., and S. F. Singer, Distribution of density in a planetary exosphere, **2**, *Phys. Fluids*, **3**, 1960a.
- Öpik, E. J., and S. F. Singer, Distribution of density in a planetary exosphere, *Phys. Fluids*, **3**, 486-488, 1960b.
- Purcell, J. D., and R. Tousey, The profile of hydrogen Lyman- $\alpha$ , *J. Geophys. Research*, **65**, 370, 1960.
- Singer, S. F., Exospheric structure deduced from satellite radiation belt and drag data, *J. Geophys. Research*, **65**, to be published, 1960.
- Yoshida, Sekiko, George H. Ludwig, and J. A. Van Allen, Distribution of trapped radiation in the geomagnetic field, *J. Geophys. Research*, **65**, 807-813, 1960.



# Whistler Dispersion and Exospheric Hydrogen Ions<sup>1</sup>

R. E. BARRINGTON AND T. NISHIZAKI

*Radio Physics Laboratory  
Defence Research Telecommunications Establishment  
Defence Research Board  
Ottawa, Canada*

**Abstract.** Accurate measurements show that the variation of whistler dispersion with frequency is of the form predicted by Storey for an exosphere consisting largely of ionized hydrogen. From these measurements the height at which protons become sufficiently numerous to affect the whistler dispersion is estimated to be 1000 km.

Whistler data are commonly reduced in terms of the 'whistler dispersion,'  $tf^{1/2}$ , which combines radio frequency  $f$  with the time  $t$  required by the frequency to traverse the transmission path from source to receiver. Measurements of whistler dispersion have been shown by Storey [1956] to be a potential source of information about the presence of hydrogen ions in the exosphere. At frequencies comparable to the gyrofrequency of an abundant positive ion of this plasma, whistler propagation is partly supported by these positive ions, causing an increase in the dispersion. Only hydrogen ions have a gyrofrequency of sufficient magnitude to cause an increase in dispersion in the frequency range of whistler propagation. The detection of this increase, which amounts at best to about 2 or 3 per cent of the total, requires high accuracy in the measurement of the arrival times of the whistler frequency components.

To achieve this accuracy, carefully selected whistlers were analyzed by a specially developed method of time-reversal tape recording and timing. Whistlers observed in Japan were used, and the hydrogen-ion effect at such low geomagnetic latitudes amounts to a greater percentage of the total dispersion than at higher latitudes. In addition, the Japanese whistlers have a much greater purity of tone than those available from other stations. In all the cases

studied, the variation of the dispersion with frequency was found to be of the form and magnitude predicted by Storey for an exosphere consisting largely of ionized hydrogen.

An attempt was made to use the observed ion effect to deduce something about the abundance and distribution of protons in the exosphere. When the motions of the light hydrogen ions are considered the total dispersion may be approximated by

$$tf^{1/2} = \frac{1}{2C} \int_{\text{path}} \frac{f_0}{f_H^{1/2}} \left\{ \frac{1}{(1 - f/f_H)^{3/2}} + \frac{1 - \epsilon[f_{iH}/(f_{iH} + f)]^2}{\left(1 - \frac{\epsilon f_{iH}}{f_{iH} + f}\right)^{1/2}} - 1 \right\} ds$$

where

$f_0$  = plasma frequency.

$f_H$  = electron gyro frequency.

$f_{iH}$  = proton gyro frequency.

$f$  = wave frequency.

$\epsilon$  = ratio of number of hydrogen ions to number of electrons.

This shows that the ion effect depends not only on the total number of protons along the whistler path but also on the fraction of the total number of ions which they constitute at any height.

It is not possible to derive a precise proton distribution from observations of whistler dispersion, but as a first approximation it was assumed that the proton and electron densities were equal down to some transition height, below

This paper was presented at the Symposium on the Exosphere and Upper  $F$  Region sponsored by Commission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.

which no protons existed. The Naval Research Laboratory's model of the variation of electron density with height was assumed and was adjusted by means of a constant factor to give the measured mean whistler dispersion. On this basis dispersion curves were calculated for various heights of the transition level and compared with the observed results. It was found that the two whistlers that were observed over a sufficiently wide frequency range to make an adequate test fitted the dispersion curves for a transition height of 1000 km appreciably better than those calculated for heights of 750 and 1250 km. Models of the electron and proton distributions of the exosphere recently published by Johnson [1960] were also used to calculate

a dispersion curve which was compared with the observed dispersions. This did not fit as well as the curves derived previously, and indicated that Johnson's model predicted too few protons in the neighborhood of 1000 to 1200 km.<sup>2</sup>

## REFERENCES

- Johnson, F. S., The ion distribution above the maximum, *J. Geophys. Research*, **65**, 5771-5774, 1960.  
Storey, L. R. O., A method to detect the presence of ionized hydrogen in the outer atmosphere, *Can. J. Phys.*, **34**, 1153-1163, 1956.

---

<sup>2</sup> Note added in proof: A detailed account of this investigation and the results of further calculations are contained in a paper now in press in the *Canadian Journal of Physics*.



# Electron Densities to 5 Earth Radii Deduced from Nose Whistlers<sup>1</sup>

R. L. SMITH AND R. A. HELLIWELL

*Radioscience Laboratory  
Stanford University, Stanford, California*

**Abstract.** Whistlers appear to propagate in columns of enhanced ionization aligned with the earth's magnetic field. From nose whistlers a value of 100 electrons/cm<sup>3</sup> is calculated at 5 earth radii.

Our study of nose whistlers has yielded new information on the structure and electron content of the outer ionosphere. The fine structure of whistlers can be attributed principally to the presence of field-aligned columns of enhanced ionization. The theory of propagation in these columns requires that the enhancement be of the order of 50 per cent of the background ionization. The lifetime of the columns is of the order of a few hours.

From each nose-whistler trace we can determine fairly accurately the effective latitude and the path of propagation along a particular line of magnetic flux. The time delay of the whistler then gives the integrated plasma frequency weighted inversely with respect to the square root of magnetic field strength.

Nose-whistler data covering the years from 1948 to 1959 have been obtained from ten different stations. They clearly indicate an annual variation of about 30 per cent in electron density in the outer ionosphere. An attempt was made to fit the data to various proposed models of the outer ionosphere. The Dungey model,  $N = N_0 \exp(2.5/R)$ , for a temperature of 1500°K is clearly inconsistent with the data. One way to modify the Dungey model to fit the data is to allow the base level of ionization to vary more than 4 to 1 over 15° of latitude (50° to 65°).

This paper was presented at the Symposium on the Exosphere and Upper *F* Region sponsored by Commission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.

Another possibility is that the temperature of the outer ionosphere is only 500°K. The same remarks apply to the model at present proposed by F. S. Johnson.

Two other models, electron density proportional to gyro frequency ( $N = Kf_H$ ), or to the product of gyro frequency and the Dungey factor ( $N = Kf_H \exp[2.5/R]$ ), fit the data fairly well. Figure 1 shows the resulting electron distribution for the gyro-frequency model.

The data and a more complete discussion of the results will be published in the near future.

**Acknowledgments.** This work was supported in part by the Air Force Office of Scientific Research under contract AF 18(603)-126.

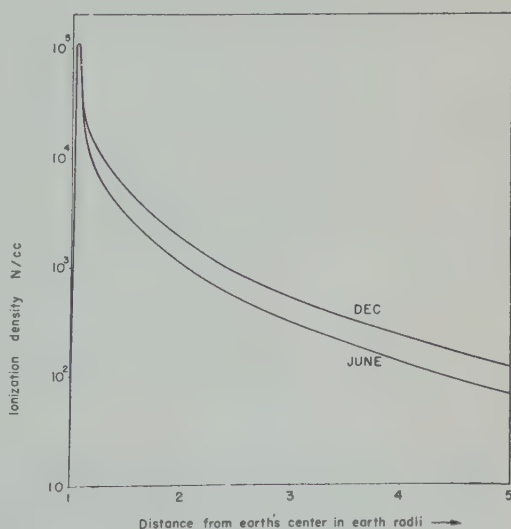


Fig. 1.





## Radio Propagation Measurements Using the Explorer VI Satellite<sup>1</sup>

CARL D. GRAVES

*Space Technology Laboratories, Inc.  
Los Angeles 45, Calif.*

**Abstract.** This paper presents the results of an experiment to measure the electron density above the ionosphere. The experiment utilized two coherent signals, one VHF and one UHF, transmitted from the Explorer VI satellite. The evidence of our data is that the electron concentration in the vicinity of the satellite, 18,000 km above the earth, was unusually high, of the order of  $10^4$  electrons/cm<sup>3</sup>. However, the results were influenced by unexpectedly high correction factors and possibly by an intense world-wide magnetic storm that occurred during the measurement period.

This paper presents the results of an experiment to measure the electron density above the ionosphere. The experiment utilized two coherent signals, one VHF (approximately 108 Mc/s) and one UHF (approximately 378 Mc/s), transmitted from the Explorer VI satellite. This satellite had a perigee of 136 miles and an apogee of 10,000 miles. Eight separate measurements were made at a receiving station in Hawaii because one of the coherent transmitters failed. To determine the electron density at the satellite, measurements were made of the Doppler difference frequency and the rate of change of the Faraday polarization rotation. Five of the recordings were of good quality and have been analyzed. Although suggestive results were obtained, enough data are available to make definite conclusions. However, these preliminary measurements indicated that the electron density appeared to be much higher than had been expected. It should be noted that the results were obtained after a large magnetic storm and also that they are contingent upon estimated corrections made for ray bending and for the effects of ionospheric static ionosphere.

When these corrections are small the simultaneous measurement of the difference between Doppler shift at two frequencies and the

change of the Faraday rotation can be used to obtain an accurate measurement of the electron density in the vicinity of the satellite. This comes about in the following way. The time rate of change of the phase of a high-frequency transmission from a satellite is

$$\frac{dP}{dt} = [\mathbf{V} \cdot \mathbf{k}]_s + \int d\mathbf{l} \cdot \frac{\partial \mathbf{k}}{\partial t} \quad (1)$$

where  $\mathbf{V}$  is the velocity of the satellite relative to the receiving station,  $\mathbf{k}$  is the propagation vector normal to the phase front with a magnitude of  $\omega n/c$ ,  $n$  is the index of refraction,  $c$  is the velocity of light in a vacuum, and  $\omega$  is the angular frequency of the transmission. The term  $\mathbf{V} \cdot \mathbf{k}$  is evaluated at the satellite, and the integral is evaluated along the path that the energy propagates. This equation holds for either polarization propagation mode. If two coherent frequencies,  $f_1$  and  $f_2 = qf_1$ , are transmitted from the satellite, the difference between them, when compared at the same frequency, that is  $f_1$  and  $qf_1$ , is

$$W = [\mathbf{V} \cdot (\mathbf{q}\mathbf{k}_1 - \mathbf{k}_2)]_s + \int d\mathbf{l} \cdot \frac{\partial}{\partial t} (\mathbf{q}\mathbf{k}_1 - \mathbf{k}_2) \quad (2)$$

This paper was presented at the Symposium on the Ionosphere and Upper  $F$  Region sponsored by Commission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.

where the difference between the two propagation paths can be neglected. The first term in equation 2, the usual *Seddon* [1953] type difference frequency term, is proportional to the electron density at the satellite. When electron-

density measurements are made in the exosphere it is necessary to include the effect of time changes in electron density in the ionosphere. In the exosphere the electron densities to be measured are small, and changes in electron density throughout the ionosphere can have an appreciable effect on the Doppler difference frequency. Therefore a part of the present experiment [Kelso, 1958] involved measuring the value of the integral appearing in equation 2. This was accomplished by recording the rate of Faraday rotation,  $\phi$ , which can be expressed using the quasi-longitudinal approximation as

$$\phi = [\mathbf{V} \cdot (\mathbf{k} - \mathbf{k}')]_s + \int dl \cdot \frac{\partial}{\partial t} (\mathbf{k} - \mathbf{k}') \quad (3)$$

where the primes distinguish one circular propagation mode from the other.<sup>2</sup>

Thus, to determine the electron density at the satellite, equations 2 and 3 are combined. But it should be noted that in equation 3 the term evaluated at the satellite is proportional to the product of the magnetic field and the electron density. Since the magnetic field at the satellite at the ranges involved in the Explorer VI measurements (12,000 to 26,000 miles above the earth) is much smaller than the magnetic field through the ionosphere, the correction to the electron density contributed by this term is negligible. Hence, at the usual high-frequency approximation for the index of refraction and quasi-longitudinal propagation [Mitra, 1952], equation 3 becomes

$$\int \frac{\partial N}{\partial t} dl = \frac{C_1 \phi}{\bar{H}_L} \quad (4)$$

where  $\bar{H}_L$  is an average value of the longitudinal component of the magnetic field through the ionosphere and  $C_1$  is a constant for any frequency. Equation 2 is most easily solved for the electron density,  $N$ , expressing  $\mathbf{V}$  in polar coordinates with respect to the receiving station; then

$$N = \frac{1}{V_R} \left[ C_2 W - \int \frac{\partial N}{\partial t} dl - C_3 V_\theta \delta \right] \quad (5)$$

where  $V_R$  and  $V_\theta$  are the radial and zenith angle polar components, respectively, of  $\mathbf{V}$ ,  $\delta$  is the angle between the propagation vectors of the two frequencies at the satellite as a result of their different ray bending. The small amount of refraction for these high frequencies allowed the neglect of higher-order terms in  $\delta$ . Combining equations 4 and 5 yields for electron density

$$N = [1/V_R][C_2 W - (C_1 \phi / \bar{H}_L) - C_3 V_\theta \delta]$$

This expression for  $N$  contains the measured quantities  $W$  and  $\phi$ , the polar components of velocity known from orbit data, and the unknown average magnetic field  $\bar{H}_L$  and ray-bending difference angle  $\delta$ . Hence, to obtain a value for the electron density it is necessary to make an estimate of  $\bar{H}_L$  and  $\delta$ . Because of the uncertainty associated with these estimates, however, it is desirable that they be small corrections to  $W$ , the predominantly determined Doppler term. This problem is inherent in the radio propagation method of determining electron densities in the exosphere; it necessitates the accumulation of considerable data at well-chosen orbit positions and times. For example, temporal changes within the ionosphere are relatively small later at night. Also, for example, the refraction correction is small when elevation angles are close to zero. Unfortunately in the short time that electron density measurements were taken from Explorer VI it was not possible to obtain data without both the correction terms were small.

Eight separate electron-density measurements were made at our Hawaiian tracking station in the period August 13 through 23, 1959, in periods varying from 20 to 70 minutes. On the latter date one of the two coherent transmitters on board the satellite failed and electron-density measurements had to be discontinued. On three occasions, August 15, 17, and 19, severe fading of the VHF signal made the data unanalyzable.<sup>3</sup> For the other five measurements, an attempt was made to estimate the ray-bending correction using ionospheric electron-density profiles obtained from the Bureau of Standards sounding station at Honolulu.

<sup>2</sup> The integral in equation 3 is evaluated along the energy propagation path of the VHF signal. In an exact treatment the two modes of propagation have slightly different propagation paths. However, to order  $1/\omega^5$  the propagation paths are identical and hence this effect can be neglected.

<sup>3</sup> The unusual fading of the VHF signal occurred almost every evening at Hawaii and will be the subject of a forthcoming publication.

a model electron-density profile<sup>4</sup> above the peak of the  $F$  layer. Unfortunately, owing to ionospheric abnormalities, electron profiles were available only for the last two days, August 22 and 23.

On August 23 the ray-bending correction was estimated to be small, approximately 10% of  $W$ , the Doppler term in equation 6. The Doppler measurement yielded an uncorrected electron density of  $10^4$  electrons/cm<sup>3</sup>. The Faraday measurement yielded an equivalent electron density due to the 'growing' ion concentration in the ionosphere of  $10^4$  electrons/cm<sup>3</sup>. This value was obtained from the measured  $\phi$  and the earth's magnetic field component along the ray path at 300 km or  $\bar{H}_L$ . The net electron density is the difference<sup>5</sup> between these values (equation 6), approximately  $2 \times 10^4$  electrons/cm<sup>3</sup>. The height of the satellite above the earth at this time was 18,000 km, and its position was 5°N latitude and 120°W longitude. This result is considerably higher than would be expected from the interpretation of whistler propagation measurements. On August 22 data were taken while the satellite was passing through apogee, and, therefore, the radial component of the satellite velocity was essentially zero. This was done to provide a check on the procedure, since under these conditions the Doppler difference frequency depends only on the ray-bending correction and the changing electron content of the ionosphere which is measured by the satellite rotation. Using the ionospheric soundings to calculate the ray-bending effect and the earth's magnetic field at 300 km with the measured Faraday rotation to calculate the ionospheric effect, results were obtained which checked with the measured value of the Doppler difference frequency to within a few percent. Even though ionospheric soundings were lacking for the other three times of

the model of the exosphere used was based on the data presented by Townsend [1959]. Note that only the absolute value of the two third terms in equation 6 is measured in this experiment. Consequently, whether these terms should be added or subtracted depends on whether the ionospheric electron content is increasing or decreasing.

measurement, an effort was made to analyze the data by calculating ray-bending corrections using a typical ionosphere. On August 16 this correction was estimated to be sufficiently small to allow a calculation of electron density. Again the data yielded electron densities of the order of  $10^4$  electrons/cm<sup>3</sup> at a height above the earth of about 20,000 km.

The evidence of our data is that the electron concentration in the vicinity of the satellite was unusually high, but in view of the small amount of data acquired and the occurrence of unexpectedly high correction factors these results can only be regarded as having order-of-magnitude validity. An event that may have influenced the results presented here was the intense worldwide magnetic storm and the abrupt increase in solar flare activity recorded on August 16 and 17. During the storm and increased activity period the low-energy scintillation counter on Explorer VI (200-kev to 1-Mev electrons) gradually increased its counting rate. By August 18 it had increased by a factor of 30, and it remained abnormally high for many days thereafter. Also, data reported on at this symposium by Farmer and Robinson indicate that electron densities have increased by an order of magnitude at heights 800 km above the earth shortly after a large solar flare. In conclusion, because of the difficulties encountered, inherent in this type of propagation measurement of electron densities in the exosphere, it seems desirable to use a more direct means for these measurements in the future.

*Acknowledgment.* I should like to thank Dr. Saul Altshuler and Dr. Richard Wagner for many helpful discussions of these measurements, and Dr. John Kelso who was a collaborator in the initial phases of the experiment.

#### REFERENCES

- Kelso, J. M., The measurement of electron densities in the outer ionosphere, *Proc. Conf. Propagation of Radio Waves, Liège, October 1958*.
- Mitra, S. K., *The Upper Atmosphere*, Asiatic Society Monograph Series, vol. 5, Calcutta, India, 2nd edition, 1952.
- Seddon, J. C., *J. Geophys. Research*, 58, 323, 1953.
- Townsend, J. W., Jr., Soviet papers presented at the rocket and satellite symposium, *Science*, 129, 80-84, 1959.





# A Sounding Rocket Measurement of Electron Densities to 1500 Kilometers<sup>1</sup>

WARREN W. BERNING

*Ballistic Research Laboratories  
Aberdeen Proving Ground, Maryland*

**Abstract.** At 0700 EST, November 10, 1959, a sounding rocket was launched from Wallops Island, Virginia, for the purpose of measuring electron densities in the high ionosphere. A continuous-wave (CW) technique was employed in the measurement, and the dispersive Doppler effects at 37 and 148 Mc/s were observed. Data are presented indicating derived electron densities to 1500 km for the up-leg of the rocket trajectory. The decrease in electron density above the  $F_2$  maximum indicates a scale height of only 30 km at 270 km but a linear scale-height gradient of 0.45 km/km from 270 to 760 km.

**Introduction.** At 0700 EST on November 10, 1959, a five-stage rocket was fired from Wallops Island, Virginia, instrumented to measure electron densities in the ionosphere. The design of the experiment and ground station, and ground station operation, were the responsibility of the Ballistic Research Laboratories. Groups at the University of Michigan, under contract, designed the nose-cone instrumentation and the rocket vehicle. The rocket was assembled and fired by personnel of the Wallops Island Facility (National Aeronautics and Space Administration) and the University of Michigan working together.

The vehicle, designated Strongarm, comprised readily available solid-propellant motors: West John, two Nike-Ajax boosters, Yardbird (air-burning Recruit), and scaled Sergeant. It was predicted that the above combination could carry 20 pounds (gross payload) to 1700 km. Actually, a peak altitude of 1800 km was attained with that payload. A high vehicle speed at relatively low altitude required teflon coating of the payload package and the final stage. To limit heat input to the nose-cone instrumentation, one temperature of the exterior surface of the Fiberglas nose (under the teflon) was recorded.

This paper was presented at the Symposium on the Exosphere and Upper  $F$  Region sponsored by Commission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.

**Design of the experiment.** The technique for the measurement of electron densities was a combination of earlier methods used by the author and by Seddon and Jackson, now of the National Aeronautics and Space Administration. Two phase-coherent, continuous-wave (CW) signals were transmitted by the rocket in flight. Frequencies of approximately 37 and 148 Mc/s were employed. Choice of the frequencies was based upon simplification of the Appleton-Hartree index-of-refraction equation and requirements imposed by size, weight, and thermal limitations in the payload carried by the rocket vehicle. Use of these frequencies resulted also in small required transmitter power outputs and simplified data reduction.

If the phase of a radio signal transmitted by a rocket in flight is represented at a ground receiver by the equation

$$\phi = \omega t - \frac{\omega}{c} \int_0^s \mu ds$$

the Doppler radian frequency, neglecting second-order effects, is given by

$$\begin{aligned} \phi_D &= \frac{\omega}{c} \left[ \frac{\partial}{\partial s} \left( \int_0^s \mu ds \right) \dot{s} + \frac{\partial}{\partial t} \int_0^s \mu ds \right] \\ &= \frac{\omega}{c} [\mu \dot{s} + s \dot{\mu}] \quad (1) \end{aligned}$$

where

$\omega$  = transmitted radian frequency.  
 $c$  = vacuum velocity of light.

$\mu$  = index of refraction, a function of  $s$  and  $t$ .  
 $s$  = path in space followed by the radio wave.

If subscripts 1 and 2 are used to designate 37- and 148-Mc/s quantities, respectively, and if radian frequencies are changed to cyclic frequencies, we may write

$$(f_{D_s} - 4f_{D_1}) \equiv B = \frac{f_2}{c} [\mu_2 V \cos k - \mu_1 V \cos (k - \Delta k) + s_2 \dot{\mu}_2 - s_1 \dot{\mu}_1]$$

or

$$B \doteq \frac{f_2 \mu_2 V \cos k}{c} \left[ 1 - \frac{\mu_1}{\mu_2} (1 + \Delta k \tan k) \right] + f_E \quad (2)$$

where

$$f_E \equiv [f_2/c][s_2 \dot{\mu}_2 - s_1 \dot{\mu}_1]$$

and use has been made of the geometry illustrated in Figure 1 and of the fact that  $\Delta k$  is a small angle at the frequencies employed. In the experiment, the quantities  $B$ ,  $f_{D_s}$ ,  $f_{D_1}$ ,  $\int_0^{s_1} \mu_1 ds$ , and  $\int_0^{s_2} \mu_2 ds$  are measured values throughout flight. The quantities  $f_1$  and  $f_2$  were measured before rocket take-off and were maintained, it is believed, to within 1 part in  $10^8$  throughout flight.

Neglecting electron collisions, absorption, ions, and magnetoionic effects, the Appleton-Hartree equation takes the simple form

$$\mu_i^2 = 1 - \frac{N_e e^2}{4\pi^2 \epsilon_0 m f_i^2} \doteq 1 - \frac{80.6 N_e}{f_i^2}$$

and, at the frequencies employed,

$$\mu_i \doteq 1 - 40.3 N_e / f_i^2 \quad (3)$$

where

$N_e$  = number density of electrons.

$e$  = electronic charge.

$m$  = electronic mass.

$\epsilon_0$  = space permittivity.

(Rationalized mksa units are used here and throughout.) Further, it will be noted that, with the exception of rocket positions very near the peak and at large zenith angles,

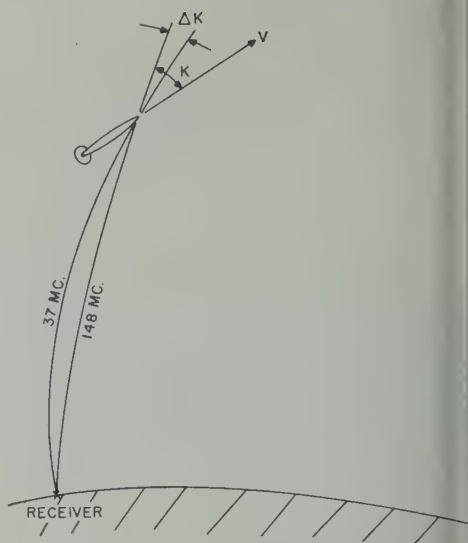


Fig. 1. Ray-path geometry.

$$f_{D_s} \doteq (f_2 \mu_2 V \cos k)/c$$

Substituting equations 3 and 4 into equation 2 and rearranging terms,

$$N_e \doteq 3.61 \times 10^{13} \left[ \frac{\pm f_{D_s} \tau + (B - f_E)}{f_{D_s} (1 \pm \tau)} \right]$$

where  $\tau \equiv \Delta k \tan k$ . The  $\pm$  sign in equation 5 simply indicates that at some point above the layer of maximum ionization the angle between the 37-Mc/s ray path and the rocket velocity vector changes from  $(k - \Delta k)$  to  $(k + \Delta k)$ . The quantity  $\tau$  in equation 5 is quite small on the ascending branch of the trajectory ( $\tau$  is for truly vertical firing) but becomes quite large on the descending branch when the ground range is larger than the height of the  $F_2$ -layer maximum. Including the  $\tau$  term requires rather detailed ray tracing if electron densities at the rocket are to be determined with accuracy on the descending branch. In ignoring this term leads to negative electron densities on the downward leg.

It is necessary now to consider the term  $f_E$  appearing in equation 5. It was noted earlier that the summed Doppler cycles, that is,

$$\frac{f}{c} \int_0^s \mu ds$$

were recorded throughout flight. The equivalent



the accumulated cycle count and the phase integral holds because integration began at zero from a known point (the launcher). Equivalence may be expressed

$$\sum_0^t n = \frac{f}{c} \int_0^s \mu ds \Big|_t = \frac{f}{c} s \bar{\mu} \Big|_t$$

At a given time, then, the accumulated Doppler for our two transmissions are

$$n_1 = \frac{f_1}{c} s_1 \bar{\mu}_1 \quad \sum n_2 = \frac{f_2}{c} s_2 \bar{\mu}_2 \quad (6)$$

Noting that  $f_2 = 4f_1$ , expressions 6 can be solved for  $s_1$  and  $s_2$ , and substituted into the expression for  $f_E$  to give

$$f_E = \left[ \frac{\dot{\bar{\mu}}_2}{\bar{\mu}_2} \sum n_2 - \frac{\dot{\bar{\mu}}_1}{\bar{\mu}_1} 4 \sum n_1 \right] \quad (7)$$

In equation 7,  $\bar{\mu}_2 \doteq \bar{\mu}_1 \doteq 1$  and  $\sum n_2 \doteq 4 \sum n_1$  (errors by less than 1 part in 100), but  $\dot{\bar{\mu}}_1 \doteq 16\dot{\bar{\mu}}_2$ . Therefore, small error is incurred when equation is rewritten

$$f_E \doteq -4\dot{\bar{\mu}}_1 \sum n_1 \quad (7')$$

It must be remembered that equation 7' requires evaluation at a specific point in space and thus a specific value of  $s$ . Accordingly, the term

$\bar{\mu}$  must include time variations in local electron density along the ray path plus the effect produced by transverse motion of the ray path in the ionosphere. Accordingly, any rigorous solution of equation 5 requires a rather complete knowledge of the ionosphere within a large volume of space as well as time variations of the ionospheric structure. Thus, the quantity  $f_E$  is usually negligibly small if the rocket is fired vertically near noon or midnight and is a maximum near sunrise or sunset with non-vertical trajectories. Referring again to equation 1, the quantity  $\mu$  was assumed to be a function of  $s$  and  $t$  only. Actually,  $\mu$  is a function of zenith angle  $\theta$  as well. Accordingly, the term  $\dot{\bar{\mu}}$  in equations 1 and 7' should be replaced by

$$\dot{\bar{\mu}} = \left( \frac{\partial \bar{\mu}}{\partial \theta} \dot{\theta} + \frac{\partial \bar{\mu}}{\partial t} \right) \Big|_{s=\text{constant}} \quad (8)$$

Fortunately, it is possible to replace the above equation for most of the trajectory by

$$\dot{\bar{\mu}} = \left( \frac{\partial \bar{\mu}}{\partial \theta} \dot{\theta} + \frac{\partial \bar{\mu}}{\partial t} \right) \Big|_{\rho=\text{constant}} \quad (8')$$

where  $\rho$  is the geometrical distance (line of sight) from the ground receiver to the rocket, and  $\theta$  the corresponding zenith angle to the rocket. With a quiet ionosphere and a time sequence

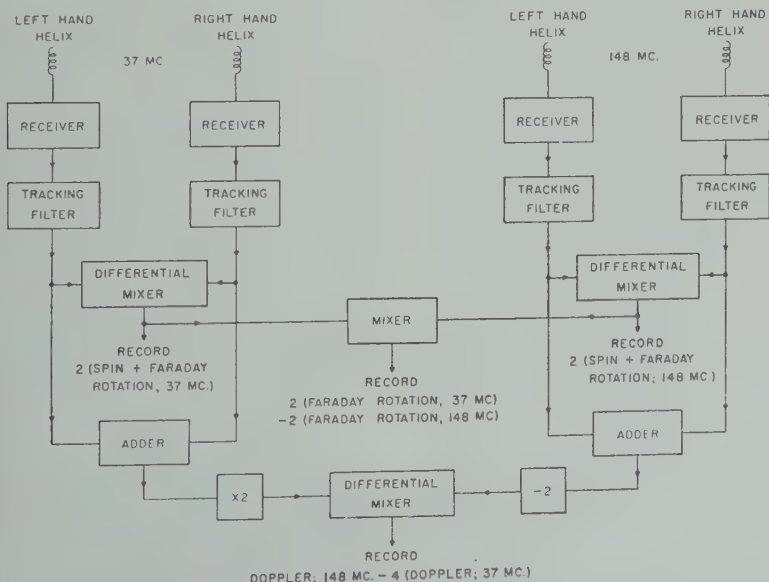


Fig. 2. Block diagram of ground-station receiving equipment.

of vertical-incidence ionograms, it is possible to estimate  $\hat{\mu}$  from equation 8 reasonably well for most of the rocket flight and thus obtain a solution for equation 5.

**Instrumental techniques.** The ground-station equipment used is shown schematically in Figure 2. Twin receiving helices of opposite polarity were used at both frequencies to permit separation of spin, Faraday effect, and dispersive Doppler and true Doppler frequencies (i.e., corrected for spin and Faraday rotation). Further, the addition of Doppler frequencies from the left-hand and right-hand helix antennas eliminates the first-order magnetoionic effect. Noting from equation 1 that the Doppler frequency is proportional to the index of refraction, an addition of Doppler frequencies will be seen to involve the sum of the ordinary and extraordinary refractive indices  $\mu_o$  and  $\mu_x$ :

$$\mu_o + \mu_x \doteq 1 - \frac{X}{2(1 + Y_L)} + 1 - \frac{X}{2(1 - Y_L)} \doteq 2\left(1 - \frac{X}{2}\right) \quad (9)$$

where

$$X = N_e e^2 / 4\pi^2 \epsilon_0 m f^2,$$

$$Y_L = \mu_o H_0 e \cos \lambda / 2\pi f m,$$

$$\mu_o = \text{space permeability.}$$

$$\lambda = \text{angle between ray path and magnetic field vector.}$$

$$H_0 = \text{magnetic field strength.}$$

At low altitudes on the down-leg of the trajectory, second-order magnetoionic effects become important because the ray paths of the ordinary and extraordinary signals become significantly different and because the quasi-longitudinal approximation expressed in equation 9 is no longer valid.

Referring again to Figure 2, attention is called to the tracking filters which follow the receiver audio outputs. These tracking filters are provided with variable bandwidths (1 to 50 cps) and permit detection and tracking of very weak signals. Through their use it was possible to maintain continuous track with transmitter power outputs of 100 and 20 mw at 37 and 148 Mc/s, respectively, despite quite inefficient rocket-borne antennas.

The phase-coherent rocket-borne transmitter

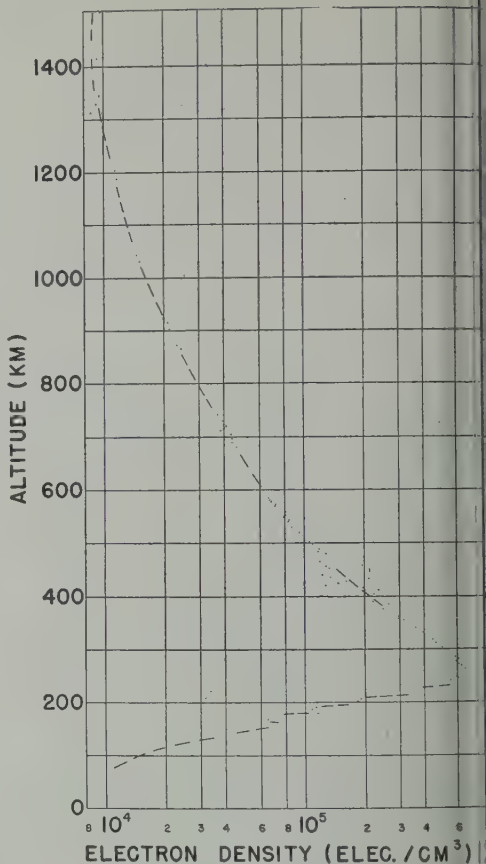


Fig. 3. Altitude profile of electron density up-leg of the trajectory.

[Cruickshank, 1960; Orr, Cath, and Darr, 1959] had a short-term frequency stability better than 1 part in 10⁸. This stability was maintained by employing a crystal oven surrounded by a partly fused bath of Cerros. This effective heat sink stabilized crystal temperature to  $\pm 0.1^\circ\text{C}$ .

**Discussion of the results.** The rocket tracked to impact, and almost all the desired data were recovered. Unfortunately, the payload tumbled or precessed so that periodically ground receivers looked at nulls in the rocket antenna patterns. Signals were not lost at these unfavorable times, but rapid changes in position resulted in scatter of the derived electron densities. For the remainder of the time, it was possible to measure the dispersive (or differential) Doppler with an accuracy of  $\pm 0.1$  cps.

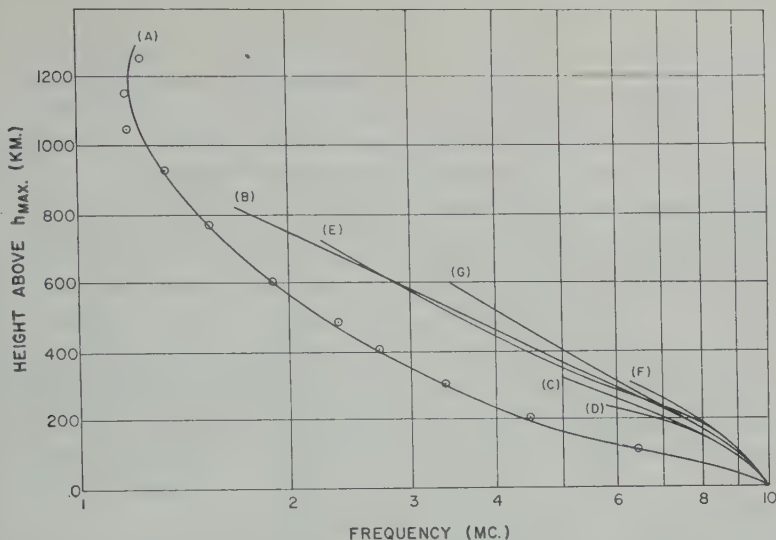


Fig. 4. Comparison of observed profile A with earlier results. Curves B, Chapman model; C and D, [unintelligible]; E 'probable profile,' Seddon; F, Farmer and Robinson; G Alpert and others [Wright, 1960]. All curves normalized to  $f_{\max} = 10$  Mc/s.

ns that errors in the derived electron density were approximately 2 per cent in the  $F_2$  layer, 15 per cent at 1000 km, and perhaps 50 per cent at 1500 km. The larger errors at high altitudes arise because changes in the ionosphere as the rocket become more important as rocket speed and local electron density decrease; these changes cannot be estimated quantitatively with great accuracy.

The derived electron density for the up-leg of the trajectory is shown in Figure 3. Relation of the data was discontinued at 1500 km (300 km short of peak altitude) because of increasing uncertainty in corrections for bending and changes in electron content as the rocket. The scatter in electron density between 400 and 500 km corresponds to a rocket position in which the transmitting antenna null is pointed toward the receiver. Similar rocket positions at higher altitudes are indicated by missing portions in the density data shown in Figure 3.

A comparison of the data derived in this experiment with earlier results is shown in Figure 4. The comparative data were obtained in a paper by Wright [1960]; reference is made to that paper for identification of the curves shown. Of particular interest is a comparison of curve B (Chapman distribution with

scale height  $H = 100$  km) with curve A (data from Fig. 3). The lowest part of curve A corresponds to a Chapman distribution with a scale height of only 30 km. However, the scale height apparently increases rapidly above 500 km. If the ionized agent above the  $F_2$  maximum is assumed to be atomic oxygen, the derived scale height of 30 km corresponds to the value for atomic oxygen. If it is further assumed that the scale height increases linearly with altitude, Nicolet [1959] has shown that the altitude variation of a constituent of the atmosphere is given by

$$n/n_0 = (g_0/g)(H/H_0)^{-(1+\beta)/\beta} \quad (10)$$

where  $n$  is the particle density and  $H$  the scale height for that constituent,  $\beta \equiv dH/dz = \text{constant}$ ,  $g$  is the gravitational acceleration (acting in the  $-z$  direction), and zero subscripts refer to some arbitrary base level. Using the data given in Figure 3, it was found that, with  $\beta = 0.45$ , Table 1 could be constructed.

The data for scale height and temperature given in Table 1 are quite different from the commonly accepted values. It must be kept in mind, however, that in deriving the information it was assumed that atomic oxygen was the principal ionized agent, that complete diffusion equilibrium existed, and that electron production



TABLE 1

| Altitude,<br>km | $n/n_0$ ,<br>eq. 10 | $n/n_0$ ,<br>Fig. 3 | $H$ , km | $T$ , °K |
|-----------------|---------------------|---------------------|----------|----------|
| 262             | 1                   | 1                   | 30       | 525      |
| 362             | 0.41                | 0.43                | 75       | 1270     |
| 462             | 0.20                | 0.19                | 120      | 1960     |
| 562             | 0.117               | 0.105               | 165      | 2640     |
| 662             | 0.074               | 0.065               | 210      | 3260     |
| 762             | 0.046               | 0.047               | 255      | 3850     |

above the  $F_2$  maximum was unimportant. If the scale heights and scale-height gradient are as shown in the table, a heat-energy flux of approximately 2 ergs/cm<sup>2</sup>/sec is implied at an altitude of 760 km.

Johnson [1960] has pointed out that, whereas the scale height appearing in the Chapman formula is that of neutral particles near the  $F_2$  maximum, the scale height for the ions well above the  $F_2$  maximum is twice as great as that for neutral particles. Thus, because the scale heights tabulated above represent the oxygen

ions, the derived temperatures would be modified above ~350 km (i.e., divided by 2) to represent the temperature of the neutral gas. With this same modification, the neutral gas scale height gradient (i.e., 0.22 km/km) implies a heat-energy flux of approximately 0.5 erg/cm<sup>2</sup>/sec.

REFERENCES

Cruickshank, W. J., Instrumentation used for ionosphere electron measurements, *Conf. Proc. 4th Natl. Convention on Military Electronics, PGMIL-IRE*, 1960.

Johnson, F. S., The exosphere and upper  $F$  region, *J. Geophys. Research*, **65**, 2571-2575, 1960.

Nicolet, M., Constitution of the atmosphere in the ionospheric levels, *J. Geophys. Research*, **64**, 2092-2101, 1959.

Orr, L. W., P. G. Cath, and B. R. Darnall, A two-frequency beacon for high-altitude ionosphere rocket research, *Univ. Mich. Final Rept. 2813-S-3-F on contract DA-20-018-509-ORD-10*, 1959.

Wright, J. W., A model of the  $F$  region above  $h_{max}F_2$ , *J. Geophys. Research*, **65**, 185-191, 1960.

# Comment on Models of the Ionosphere above $h_{\max}F_2^1$

J. W. WRIGHT

National Bureau of Standards  
Boulder, Colorado

**Abstract.** Evidence for a gradient of scale height in the  $F$  region is shown, and discussed in relation to a simple Chapman model of the  $F$  region above  $h_{\max}F_2$ . It is suggested that a similar model, but allowing for a scale-height gradient, may give somewhat better agreement with recent observations.

The model proposed in a recent paper [Wright, 1960] showed rather good agreement with other data, provided one used a scale height of 100 km (for the ionizable material) above  $h_{\max}F_2$ . It is important to realize that, using this model, the ratio  $N_{\text{above}}/N_{\text{below}}$  is directly proportional to this scale height. On the basis of the theoretical work of Dungey [1956], and the expectation that diffusion and attachment-like loss hold the layer in at least pseudo-Chapman form near the peak, we find that the peak should then be very nearly parabolic, as given by

$$N = N_{\max} \{1 - [(h - h_{\max})/2H]^2\}$$

From our  $N(h)$  profiles, we have for some time been computing the above quantity  $H$  as the quarter-thickness of the parabola best fitting the observed (bottomside) peak of the  $F$  layer. This quantity should then nearly approximate the scale height near the level  $h_{\max}$ .

In Figure 1 some data for this quantity (which we call SCAT) are shown. For the months March, April, May 1959 at Puerto Rico, and for the 3 hours centered on midnight, SCAT has been plotted against  $h_{\max}$ . Several things may be seen:

(a) The range of SCAT is between 30 and 80 km. These values are in quite reasonable agreement with estimates of the scale height for atomic oxygen at  $F$ -region heights.

(b) SCAT is seen to vary, in general, linearly with  $h_{\max}$ . A gradient  $dH/dh$  of about 0.2 would be deduced from this figure. This may be a

reasonable value for the scale-height gradient in the upper  $F$  region.

Remembering the model discussed above, it is not reasonable to expect that  $H$  in the  $F$  region changes abruptly from a value of, say, 50 km below  $h_{\max}$  to a value of 100 km above. It may be shown that, in the presence of a scale-height gradient, and still considering diffusion and attachment-like loss, a distribution very much like Chapman's formula again results.

A model will be devised on this basis; it is expected to have the following properties:

(a) The scale-height gradient will be the only arbitrarily chosen parameter.

(b) The scale height at the peak will be determined by the quarter-thickness of the layer just below the peak; the model will then be fitted to  $h_{\max}$ , and  $H$  will be determined from ground-based soundings.

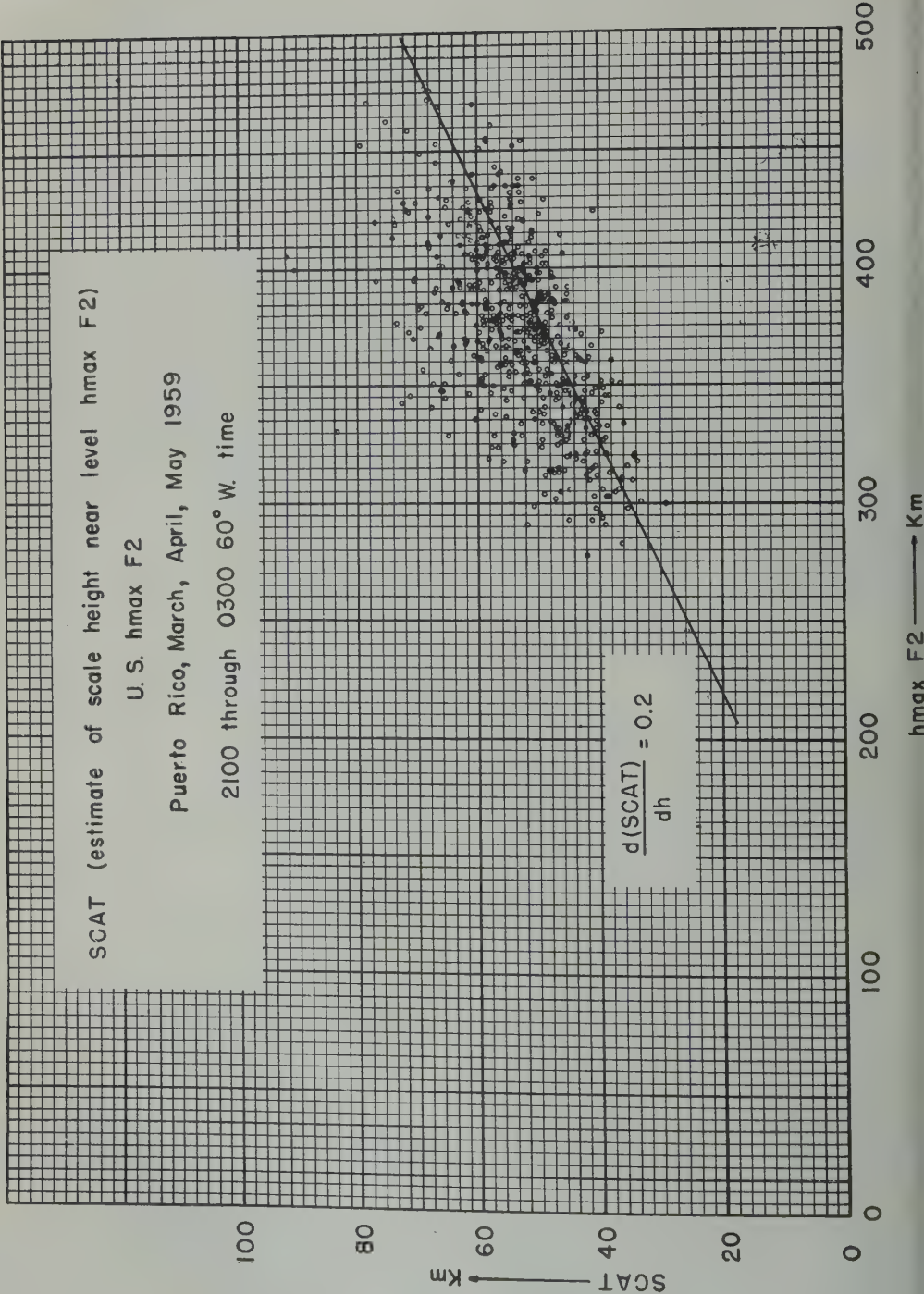
(c) The thickness of the model immediately above the peak will be only slightly larger than that below, in agreement with some of the recent rocket and satellite profiles [Berning, 1960; Garriott, 1960].

(d) The ratio of electron content above the peak to that below will be smaller than the values of 3.5-6 given by the constant 100-km scale height of the older model; this would appear also to be in agreement with recent results.

## REFERENCES

- Berning, W. W., A sounding rocket measurement of electron densities to 1500 kilometers, *J. Geophys. Research*, **65**, 2589-2594, 1960.
- Dungey, J. W., The effect of ambipolar diffusion in the night-time  $F$  layer, *J. Atmospheric and Terrest. Phys.*, **9**, 90-102, 1956.
- Garriott, O. K., The determination of ionospheric electron content and distribution from satellite observations, Part 2, *J. Geophys. Research*, **65**, 1151-1157, 1960.
- Wright, J. W., A model of the  $F$  region above  $h_{\max}F_2$ , *J. Geophys. Research*, **65**, 185-191, 1960.

<sup>1</sup>This paper was presented at the Symposium on the Exosphere and Upper  $F$  Region sponsored by Commission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.





# Electron-Density Distribution in the Upper Ionosphere from Rocket Measurements<sup>1</sup>

JOHN S. NISBET

*Ionosphere Research Laboratory  
The Pennsylvania State University  
University Park, Pennsylvania*

**Abstract.** Measurements are given of electron-density distributions above the maximum of the *F* layer from seven rocket flights. The missiles used for these experiments had large horizontal velocity components, and a matrix method was employed to derive the profiles, account being taken of horizontal gradients, refraction, and the varying ray path zenith angles in the ionosphere. Two methods of measuring ionospheric dispersion were compared: Faraday rotation, and a method using the difference between the apparent positions of the missile obtained by two tracking systems operating on different frequencies. Faraday rotation measurements at two receiving stations and on two frequencies were compared to check the horizontal gradient program and the polarization of the rocket antenna. The implications of the results are discussed in terms of the structure of the layer and its diurnal variation.

Figure 1 shows electron-density measurements during seven rocket flights. Faraday rotation measurements were made on all flights, and multiple frequency dispersion measurements on three of them.

The program employed for the calculations took account of variations in the local ray zenith angle with altitude and range time, variations in the value of the earth's magnetic field along the ray path, and horizontal gradients in electron density.

Only seven rocket results are available, obtained at different times of day, year, and sunspot cycle, any conclusions about the diurnal variation of parameters must necessarily be tentative.

The electron content of a vertical column of one square meter cross section between 400 and 500 km altitude is shown in Figure 2 as a function of local time. The length of the lines is an indication of the accuracy of the measurement. A Russian measurement by Gringauz [1958] has also been included. A large diurnal variation is indicated at these altitudes, with a maximum

<sup>1</sup>This paper was presented at the Symposium on the Ionosphere and Upper *F* Region sponsored by the Commission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.

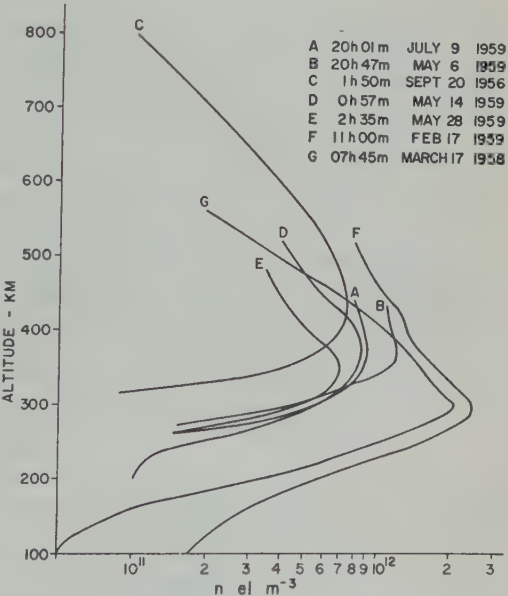


Fig. 1. Rocket electron-density measurements.

at around 16 hours. The electron density at the maximum of the layer appears to have a somewhat different variation.

Of considerable interest to those working with true-height scaling of ionograms is the accuracy with which the height of the maximum can be determined. In our rocket experiments this parameter could be measured from the profile

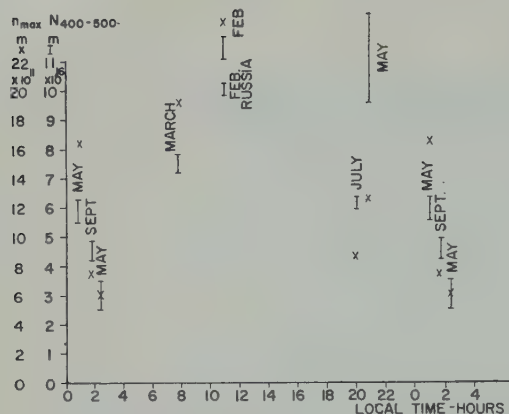


Fig. 2. Electron content from 400 to 500 km as a function of local time.

for comparison purposes. During the day the agreement was quite good, but at night, even when extreme care was taken in the reduction, the four most accurate measurements showed the altitude of the maxima to be 20 to 40 km higher than that given by the ionograms, though the agreement of the profiles just below the maximum was considerably better than this. Such an error in the height of the maximum would result in a considerable error in the electron content integrated up to the height of the maximum and in the scale heights indicated by satellite electron-content measurements made up to altitudes not greatly in excess of the maximum. It is suggested that satellite and moon-echo measurements of electron content be related to electron-content measurements up to a fixed altitude instead of up to the maximum. The

measurements should be more accurate, and diurnal variations in electron content above fixed altitude should be more easily interpreted in terms of transport mechanisms and processes in the upper ionosphere.

Figure 3 shows the seven rocket measurements of electron density in a normalized form compared with those of others. The measurements are listed in order of decreasing percentage vertical gradient. Curves A and B were made shortly after ground sunset, and the electron density appears to decrease very slowly up to 150 km above the maximum. Curve G represents a measurement made after sunrise and shows the largest percentage gradient. Between these extreme values lie the day and the night profiles.

It is postulated that the diurnal variation shown is due, in part, to the diffusion process causing the relative electron density to lag the diffusion equilibrium value.

It is apparent that these results are not consistent with a simple Chapman model of constant scale height. If a model of the neutral atmosphere is assumed, the normalized electron density height profiles can be plotted in terms of number of scale heights above the maximum. This has been done in Figure 4, using a model of the neutral atmosphere [Kallin, 1959]. Rocket results are shown by points. It is apparent that the spread in the results at different times of day has been reduced to 10 per cent at two scale heights above the maximum compared with  $\pm 39$  per cent at 150 km above the maximum in Figure 3. The solid line is the Chapman function shown in

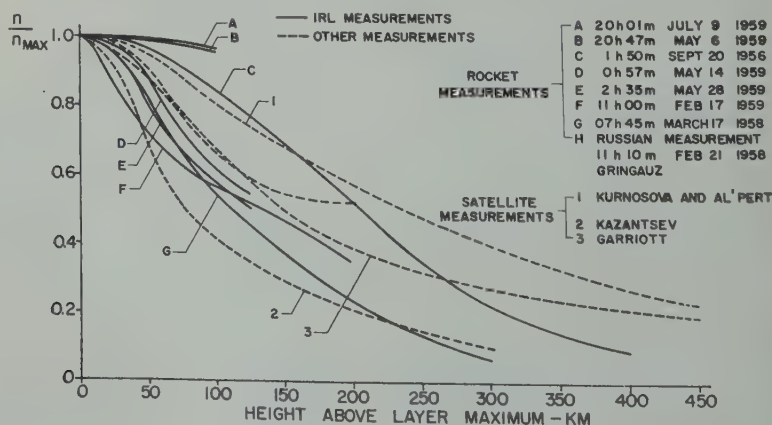


Fig. 3. Comparison of normalized electron-density profiles from satellite measurements with rocket results.

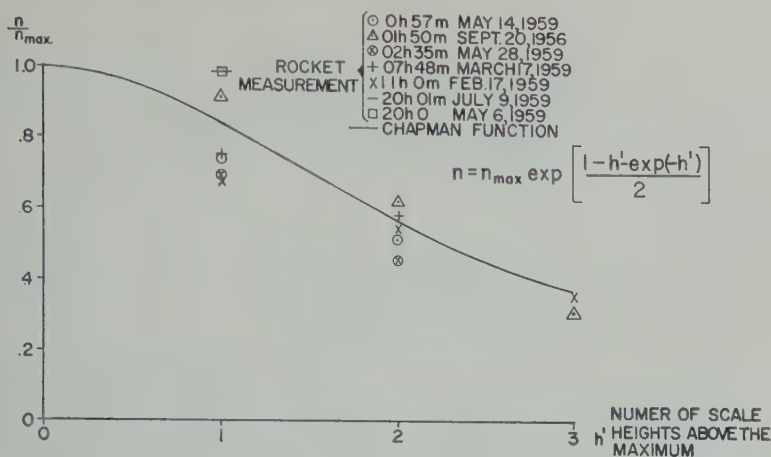


Fig. 4. Electron-density variation with number of scale heights above the maximum.

the plot is in terms of number of scale heights above the maximum, it has unique properties. It is apparent that this curve agrees well with the mean value of the experimental points.

A more complete description of the analysis and conclusions is now in preparation and will be submitted for publication shortly.

**Acknowledgment.** The research reported in this paper is carried out under the sponsorship of the Ballistic Missile Agency under contract DA-11-ORD-577.

#### REFERENCES

Amuz, K. I., Rocket measurements of the electron concentration in the ionosphere using

an ultra-shortwave dispersion interferometer, *Doklady Akad. Nauk SSSR*, 1206, 1234-1237, 1958.

Garriott, O. K., The determination of ionospheric electron content and distribution from satellite observations Part 2: Results of Analysis, *J. Geophys. Research*, 65, 1151-1157, 1960.

Kallmann, H. K., A preliminary model atmosphere based on rocket and satellite data, *J. Geophys. Research*, 64, 615-623, 1959.

Kazantsev, A. N., Absorption and electron distribution in the  $F_2$  Layer determined from measurements of transmitted signals from earth satellites, *Planet. Space Sci.*, 1, 130-135, 1959.

Kurnosova, L. V., and Ya. L. Al'pert, Certain results of research performed using artificial earth satellites, *Priroda*, 6, 85, 1958 *Translation Physics Express* 1, 32, Sept. 1958.





# The Determination of Ionospheric Electron Content from Satellite Doppler Measurements

## 1. Method of Analysis<sup>1</sup>

W. J. Ross

*Ionosphere Research Laboratory  
The Pennsylvania State University  
University Park, Pennsylvania*

**Abstract.** A procedure for determining the ionospheric electron content up to the height of an active satellite from Doppler data is developed. The equations derived from first-order theory are discussed and corrected separately for earth curvature, large refraction, off-zenith orbit, vertical satellite motion, horizontal ionospheric variations, and the effects of the earth's magnetic field. The methods were developed initially for use with the harmonic radiations from satellite 1958<sub>2</sub> at frequencies of approximately 20 and 40 Mc/s, but may be adapted to other harmonic frequency ranges. The results obtained by applying this method to the experimental data from 1958<sub>2</sub> are presented in part 2.

**Introduction.** The measurement of propagation effects in the radio signals received from a satellite transmitter affords opportunity for the determination of ionospheric properties. In particular, if the satellite orbit lies above the critical frequency of maximum electron density it is possible to derive information about the upper ionosphere.

A ground-based observer measuring the frequency of a radio wave from a satellite will find that it differs from the transmitted frequency by the Doppler frequency given by the rate of change of the phase path length in wavelengths of the radio wave between the transmitter and receiver. Owing to ionospheric refraction, the phase path length will be different from its free-space value, and, furthermore, since the ionospheric refractive index is dispersive, the relative change in length of the free-space length will depend on frequency.

A direct evaluation of the ionospheric effect on the Doppler frequency is often not possible, as it would require accurate knowledge of the position of the satellite with time and of its radiated

frequency. The frequency-dependent part of the ionospheric Doppler effect, however, can be measured accurately, without recourse to orbital information, provided that two satellite transmitters radiate coherently at different frequencies. For example, if harmonically related satellite transmitters are available, the departure from a harmonic relationship of the received signals can be measured and attributed to the dispersive properties of the changing ionospheric path of propagation.

Interpretation of these 'dispersive Doppler' frequency patterns will require knowledge of the orbital parameters, but the accuracy to which this information is required is much lower than that for direct measurement of the ionospheric Doppler effect.

This paper will deal with the analysis of the dispersive Doppler effect; other propagation effects that may also yield information about the ionosphere will not be considered. The methods to be described were developed initially for use with the radio signals from satellite 1958<sub>2</sub>, which radiated keyed signals at frequencies of approximately 20.0047 Mc/s and at harmonics of this frequency. With suitable modification the methods are applicable to other frequency ranges.

The paper will consist of two parts, the first being a development of the methods of data

<sup>1</sup>This paper was presented at the Symposium on the Exosphere and Upper F Region sponsored by Commission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.

analysis and the second a presentation and discussion of experimental results.

2. *Outline of the problem.* The exact solution of the propagation equations for paths between a satellite and a ground-base observer is of considerable difficulty except in the simplest cases. The inverse problem of the interpretation of propagation effects in terms of ionospheric properties is even more difficult and is probably not solvable in the general case. It therefore becomes of great importance to derive approximate solutions to these problems so that the principal features of the effects can be evaluated. Expressions for the propagation effects produced by simply assumed models of the ionosphere will first be derived and then inverted to permit the data to be interpreted in terms of the assumed models. The effects of model deviations on the interpretation of data will be considered.

The simple case of a satellite moving horizontally at observer zenith over a plane earth and horizontally stratified ionosphere system in which the refraction is small will be treated, and the dispersive Doppler frequency will be related to the vertical columnar electron density up to the height of the satellite. The following deviations from this initial model will be analyzed, and correction factors for the effects of each of these perturbations will be calculated and applied independently: (1) spherical-earth considerations; (2) large refraction conditions; (3) off-zenith passage; (4) vertical satellite motion; (5) horizontal ionospheric variations; (6) magnetic-field effects.

The Doppler effect problem has been analyzed by other workers [Aitchison and Weekes, 1959; Hibberd and Thomas, 1959]. These analyses, however, do not consider all the corrections to the simple theory found necessary for evaluation of the data to be presented here and are not in the most useful form for the inclusion of orbital and other data available to us. Consequently, the methods to be discussed here, which were developed independently, will be presented in their complete form.

3. *First-order theory.* Weekes [1958] has shown that, for a satellite moving in the vertical plane through an observer on the earth's surface beneath a spherically stratified isotropic ionosphere, the rate of change of phase path length,  $P$ , between satellite and observer is given

by

$$\dot{P} = c \frac{\Delta f}{f} = -r_0 \dot{\theta}_s \sin i - r_0 (q^2 - \sin^2 i)^{1/2} \dot{r}_s / r_s \quad (3)$$

and the geocentric angular separation of satellite and observer,  $\theta_s$ , by

$$\theta_s = \int_{r_0}^{r_s} \frac{\sin i \, dr}{(q^2 - \sin^2 i)^{1/2} r} \quad (3)$$

where

$r_0$  is the earth radius.

$r_s$  is the radial distance from the earth's center to the satellite.

$i$  is the angle of incidence at the observer; the ray from the satellite.

$q = r\mu/r_0$ .

$\mu$  is the refractive index at distance  $r$  from the earth's center.

$s$  denotes the value at the satellite.

For the case of horizontal satellite motion over the spherical earth-ionosphere system, equation 3.1 yields

$$\dot{P} = -r_0 \dot{\theta}_s \sin i \quad (3)$$

The dispersive Doppler frequency between two transmitted frequencies  $f_1, f_2$ , will be related to the difference in  $\dot{P}$  at the two frequencies. It becomes convenient to define a quantity

$$R = (\ddot{P}_2 - \ddot{P}_1) / \ddot{P}_1 \quad (4)$$

which is the ratio of rates of change of dispersive phase path velocity and of phase path velocity at one of the frequencies.

For our spherically symmetrical model at zenith

$$R = \frac{\int_{r_0}^{r_s} \frac{dr}{\mu_1 r^2} - \int_{r_0}^{r_s} \frac{dr}{\mu_2 r^2}}{\int_{r_0}^{r_s} \frac{dr}{\mu_2 r^2}} \quad (4)$$

For the special case of a plane earth,  $f_2 = 2f_1$  and assuming the Eccles-Larmor form of the refractive index,

$$\mu_1^2 = 1 - \alpha N \quad (4)$$

and

$$\mu_2^2 = 1 - \alpha N/4 \quad (4)$$



re  $N$  is the electron density at a given height, and assuming also  $\alpha N \ll 1$ , equation results in

$$R \simeq \frac{3}{8} \alpha \bar{N} \quad (3.8)$$

$$\bar{N} = \frac{1}{h} \int_{h_0}^{h_s} N dh$$

$$\therefore \alpha \bar{N} = 8R/3 \quad (3.9)$$

he quantity  $R$ , which can be determined solely from radio observation, is proportional to the mean electron density up to the height of the satellite in this simple case.

Similar results have been derived previously by Weekes [1958] and others.

*Spherical-earth corrections.* For the case of a spherically stratified ionosphere in which  $\alpha N \ll 1$ , and for  $f_2 = 2f_1$ , equation 3.5 becomes

$$R = \frac{\frac{3}{8} \alpha \int_{r_0}^{r_s} \frac{N}{r^2} dr}{\int_{r_0}^{r_s} \frac{dr}{r^2}} = \frac{3}{8} \alpha \frac{r_0 r_s}{r_s - r_0} \cdot \int_{r_0}^{r_s} \frac{N}{r^2} dr \quad (4.1)$$

If the satellite is in orbit above most of the ionization, to a first approximation the ionization can be considered to be concentrated near the level of maximum density  $r_m$ . Then

$$R = \frac{3}{8} \alpha \bar{N} (r_0 r_s / r_m^2)$$

$$\alpha \bar{N} = (8R/3) (r_m^2 / r_0 r_s) \quad (4.2)$$

The effects of layer curvature are contained in the last factor, which corrects for the heavier weighting given to lower ionization. For moderate satellite altitudes (<2000 km, say) this factor can be written approximately as

$$1 + (2h_m - h_s)/r_0$$

considering  $h_m \simeq 400$  km,  $r_0 = 6373$  km, as

$$1 + (800 - h_s)/6373 \quad (4.3)$$

The corrections associated with this effect are generally of the order of a few per cent only and may be either positive or negative as  $h_s$  approaches 800 km.

5. *Large refraction effects.* For the plane-earth model at zenith the binomial expansion of equation 3.6 in the form

$$1/\mu_1 = 1 + \frac{1}{2} \alpha N$$

will introduce appreciable error if  $\alpha N$  is not small.

Further, incipient ray bending effects are appreciable even at zenith, as can readily be seen from comparison of the exact expression (from equation 3.5)

$$R + 1 = \frac{\int_{h_0}^{h_s} \frac{1}{\mu_1} dh}{\int_{h_0}^{h_s} \frac{1}{\mu_2} dh} \quad (5.1)$$

with the 'straight-line approximation' equivalent expression

$$R + 1 = \frac{\int_{h_0}^{h_s} \mu_2 dh}{\int_{h_0}^{h_s} \mu_1 dh}$$

Binomial expansion of  $1/\mu_1$  and  $1/\mu_2$  in equation 5.1 and termwise integration yield the result

$$R + 1 = \frac{1 + \frac{1}{2} \alpha \bar{N} + \frac{3}{8} \alpha^2 \bar{N}^2 + \frac{5}{16} \alpha^3 \bar{N}^3 + \dots}{1 + \frac{1}{8} \alpha \bar{N} + \frac{3}{128} \alpha^2 \bar{N}^2 + \frac{5}{1024} \alpha^3 \bar{N}^3 + \dots}$$

whence

$$\alpha \bar{N} \simeq \frac{8R}{3} \frac{1 + \frac{1}{8} \alpha \bar{N} + \frac{3}{128} \alpha^2 \bar{N}^2 + \dots}{1 + \frac{15}{16} \frac{\alpha \bar{N}^2}{\bar{N}} + \frac{105}{128} \frac{\alpha^2 \bar{N}^3}{\bar{N}} + \dots}$$

Since the relationship between the mean values of powers of  $N$  is not known, some estimate must be made of the correction to be applied to the simple theory. It has been found from the exact analysis of a number of assumed ionization profiles that the mean electron density can be represented approximately by

$$\alpha \bar{N} \simeq \frac{8R}{3} \left( 1 + \frac{R}{5} \right) (1 - F \alpha N_e) \quad (5.2)$$

where  $N_e$  is the maximum electron density of the model and  $F$  is a form factor for the model under consideration.

The functional form of this expression follows from equation 5.1, the second factor representing the numerator and the last the denominator.

The value of  $F$  has a maximum of about unity for a slab layer of constant density, and is a measure of the effective fraction of the ionization at maximum density. The form of the actual ionosphere is not known above the level of maximum density but for preliminary analysis purposes may be considered to be given approximately by the form of the Chapman function

$$N = N_c \exp[1 - z - e^{-z}] \quad (5.3)$$

where  $z$  is measured from the level of maximum density. When the satellite orbit lies far above the level  $z = 0$ , the value of  $F$  is found to be 0.66. This value of  $F$  has been taken for purposes of initial data reduction. It may be altered later on the basis of these or other investigations, and, as will become evident in the second part of this paper, it may have diurnal and seasonal dependence also.

The most extreme refraction conditions occur at temperate latitudes near winter noon when  $\alpha N_c$  may be as large as 0.65 for a 20-Mc/s signal. The large refraction correction factor will then have the value of about 0.57, so that any error in the assumed value of  $F$  will appear to about the same extent in  $\bar{N}$ . Generally, however, errors in  $F$  will not affect  $\bar{N}$  greatly.

6. *Off-zenith passage.* In general the satellite will not pass through the observer's zenith but will approach to some minimum zenith angle,  $i_{\min}$ . The Doppler frequency shift will still be governed [Weekes, 1958] by the rate of change of zenith angle, if horizontal symmetry and horizontal motion are assumed. The value of  $R$  for a given ionosphere will differ from its zenith value, owing to the differential ray bending between the propagation paths at the two frequencies. Because the variation of  $R$  is closely linked with ray bending it was considered necessary to examine the effect under large refraction conditions.

Exact solutions of the propagation paths for zenith angles of up to  $40^\circ$  were obtained by iterative methods for a variety of slab models of the ionosphere of varying thickness and critical frequency. Comparison of the values of  $R$  so obtained with the zenith values yielded correction factors which were fitted with an expression of the form  $(1 - G \sin^2 i_{\min})$ .

The evaluation of  $G$  in terms of model parameters showed that it was largely controlled by the critical frequency of the assumed slab model and only slightly dependent on the thickness of the slab for reasonable ranges of thickness.

A representative slab model was chosen in which the ionospheric thickness was 0.25 of the satellite height. This would correspond to a 250-km slab for a satellite altitude of 1000 km and it will be seen in part 2 that these values are quite representative of the actual conditions.

A value of

$$G = 2(1 - \mu_c)^{1.25}$$

where  $\mu_c$  is the refractive index of the slab, was found to give a good fit to these data and has been used to correct the experimental data.

The adequacy of slab-model calculations is believed to be quite good, as the ray bending occurs principally near the nose of the  $F_2$  layer where electron density is almost constant with height. However, records in which the correction exceeds 20 per cent or in which  $i_{\min}$  exceeds  $45^\circ$  have not been included in the later tabulations of experimental data.

7. *Vertical satellite motion.* The above considerations apply to a satellite moving horizontally. When a vertical component of motion is also present the system loses its symmetry and the dispersive Doppler frequency zero crossing longer occurs at the closest approach to zenith but some distance in the direction of decreasing satellite height.

The simplest approach to the evaluation of the effects of vertical motion is to assume condition  $\alpha N \ll 1$  to hold, so that phase path lengths can be approximated by straight-line integration between transmitter and receiver rather than along the ray path. Then

$$P_2 - P_1 \doteq \frac{3}{8} \alpha \cdot \int N dr$$

and

$$\dot{P} \cdot R \doteq \frac{3}{8} \alpha \frac{d^2}{dt^2} \left[ \int N dr \right] \quad (7.1)$$

Evaluation of equation 7.1 for the plane-earth case at zenith yields approximately

$$R \simeq \frac{3}{8} \alpha \bar{N} \left[ 1 + \frac{h_p^2}{h_p \theta^2 H} \cdot \frac{N_c}{\bar{N}} + \frac{\ddot{h}_p}{h_p \theta^2} \left( \frac{N_c}{\bar{N}} - 1 \right) \right] \quad (7.2)$$

The  $\bar{N}$  is the mean electron density to the satellite at zenith and  $H$  is the characteristic length of the assumed exponential decrease of electron density with height.

Examination of this expression under the conditions of vertical motion of 1958  $\delta_2$ , pertaining to the data discussed in part 2 showed that,  $\delta_2$  does not exceed  $\bar{N}$  and if  $H$  is taken to be 100 km, the corrections in  $R$  to first-order theory are never in excess of 5 per cent and are principally due to vertical motion through a varying ionosphere rather than to vertical acceleration. It follows, therefore, that, if the analysis is restricted to a satellite passes well above the ionosphere, especially to those with small vertical velocity, a large correction to first-order theory is unnecessary.

**Horizontal ionospheric gradients.** It is well known that the electron density of the ionosphere varies from place to place, the principal variations being due to diurnal changes and latitude effects, although some smaller-scale irregularities may be present also.

We assume, as in the previous section, that straight-line integration of phase paths is permissible and if we also consider the satellite as moving horizontally well above a plane ionosphere which can to a first approximation be regarded as localized at a fixed height  $h_m$  to have an electron content which has no lateral dependence, then it follows that near

$$\frac{3}{8} \alpha \bar{N} \left[ 1 + h_m^2 \frac{1}{\int N dh} \frac{\partial^2}{\partial x^2} \left( \int N dh \right) \right] \quad (8.1)$$

where  $x$  is measured in the direction of satellite motion.

an exponential lateral dependence of electron content is assumed, of characteristic length  $L$ ,

$$R = \frac{3}{8} \alpha \bar{N} [1 + h_m^2 / L^2] \quad (8.2)$$

Thus if the scale of the irregularities is of the order of several thousand kilometers the corrections to be applied are quite small. The effect will probably be greatest near sunrise, when the vertical gradients in the ionosphere may vary rapidly with distance, but a reliable estimate of correction factors cannot readily be made with the present available information.

Small-scale irregularities will produce scatter in the dispersive Doppler patterns but should not affect the measurement of total electron content other than to increase the reading errors.

**9. Magnetic-field effects.** The earth's magnetic field causes the ionosphere to be anisotropic and birefringent. When, as in the present case, the wave frequency is much greater than the gyro frequency of the ionospheric electrons and is also considerably greater than the plasma frequency, and wave propagation is not closely normal to the magnetic field, the quasi-longitudinal (QL) form of the Appleton-Hartree index is applicable [Ratcliffe, 1959] and is given approximately by

$$\mu_1^2 = 1 - \alpha N / (1 \pm f_L / f_1) \quad (9.1)$$

where  $f_L$  is the gyro frequency of electrons in the component of the magnetic field resolved along the direction of propagation.

The upper and lower signs refer to the extraordinary and ordinary modes of propagation respectively.

When  $\mu$  is close to unity and  $f = 20$  Mc/s, at zenith at the latitude of University Park, Pennsylvania (dip angle  $72^\circ$ ), the anisotropy in  $(1 - \mu)$  is about 6 per cent at  $F$ -layer heights.

This anisotropy causes a deviation of the ray from the phase path calculated using the isotropic index of equation 3.6. However, near zenith when the value of  $f_L$  changes very slowly with zenith angle, and considering that integration along paths close to the ray path will introduce only second-order errors in phase path length, it is a good approximation to consider the phase path length to be given by the integral along the isotropic phase path. Similar conclusions were reached by Garriott [1960] in analyzing magneto-ionic ray splitting as it affected polarization rotation studies.

If one observes one or other mode of propagation at each of two initially harmonic frequencies, the value of  $\alpha$  must be adjusted accordingly by approximately 6 per cent, being larger for ordinary mode observations. However, if linearly polarized receiving antennas are used and readings are taken on a linearly polarized source near times of maximum signal when the two magneto-ionic modes are in phase in the receiving antennas, the measured dispersive Doppler frequency will be approximately the mean of the separate dispersive Doppler fre-



quencies of the two modes, and to a good approximation the effect of the magnetic field may be neglected.

10. *Summary.* Combining all the considerations of sections 1 through 9, the resultant form of the equation for  $\alpha\bar{N}$  becomes

$$\alpha\bar{N} = \frac{8R}{3} \left( \frac{1+R}{5} \right) \cdot (1 - 0.66\alpha N_e) \cdot [1 - 2(1 - \mu_e)^{1.25} \sin^2 i_{\min}] \cdot \left[ 1 + \frac{800 - h_s}{6373} \right] \quad (10.1)$$

The corrections for vertical motion (section 7) and horizontal ionospheric gradients (section 8) have not been included because they cannot be estimated reliably.

The quantities  $h_s$  and  $i_{\min}$  can be obtained directly from orbital data,  $N_e$  and  $\mu_e$  from ionosonde data;  $R$  is an experimentally measured quantity. It has been found more convenient and accurate, however, to evaluate  $R$  partly on the basis of orbital data.

The over-all accuracy of the above expression when applied to experimental data cannot be tested directly, but from the earlier discussion it is believed that the error in  $\bar{N}$  and hence in the total ionospheric content,  $N_T$ , to the height of the satellite, does not exceed 10 per cent for satellite passes lying above 800-km altitude, have fairly small eccentricity ( $<0.1$ ), pass within about  $30^\circ$  of zenith, and occur at times for which refraction is not too great ( $f_oF_2 < 14$  Mc/s). Indirect evidence will be shown in part 2 for believing that the actual accuracy is rather better than 10 per cent.

Occasionally it is found that the nonuni-

formities of the ionosphere itself produce a that cannot be scaled to the above accuracy and on almost all occasions an uncertainty of a few per cent occurs due to the variation of the ionosphere as the ray path moves through it.

Even with these uncertainties, the above method will provide a basis by which the electron content of the ionosphere can be measured and its dependence on various controlling influences determined.

*Acknowledgments.* Discussions of this work with Staff members of the Ionosphere Research Laboratory, especially Dr. J. S. Nisbet, have been very helpful and are greatly appreciated.

The work has been supported by the National Science Foundation under grant Y32.44/270 and more recently by the Wright Air Development Division and the National Aeronautics and Space Administration under contract AF33(616)-611.

#### REFERENCES

- Aitchison, G. J., and K. Weekes, Some deductions of ionospheric information from the observations of emissions from satellite 1957 $\alpha_2$  1, theory of the analysis, *J. Atmospheric and Terrest. Phys.*, 14, 236-243, 1959.
- Garriott, O. K., The determination of ionospheric electron content and distribution from satellite observations, 1, Theory of the analysis, *Geophys. Research*, 65, 1139-1150, 1960.
- Hibberd, F. H., and J. A. Thomas, The determination of the electron distribution in the upper ionosphere from satellite Doppler observations, *J. Atmospheric and Terrest. Phys.*, 17, 7-19, 1959.
- Ratcliffe, J. A., *The Magneto-ionic Theory and Its Applications to the Ionosphere*, chapter 1, Cambridge University Press, 1959.
- Weekes, K., On the interpretation of the Doppler effect from senders in an artificial satellite, *J. Atmospheric and Terrest. Phys.*, 12, 335-341, 1958.

# The Determination of Ionospheric Electron Content from Satellite Doppler Measurements

## 2. Experimental Results<sup>1</sup>

W. J. Ross

*Ionosphere Research Laboratory  
The Pennsylvania State University  
University Park, Pennsylvania*

**Abstract.** The preliminary results of the analysis of daytime Doppler frequency records from satellite 1958 $\delta_2$  for the period September 1958 to December 1959 are presented in the form of ionospheric electron content derived using the methods of part 1. Strong seasonal effects are found, the content being lower and the equivalent slab thickness greater in summer than in winter. Magnetic control of the summer ionosphere content is indicated, although the equivalent slab thickness is almost constant. The values of content obtained agree fairly well with the results of independent polarization rotation studies.

**Introduction.** Part 1 developed a method of analysis whereby the experimentally measured dispersive Doppler frequency from an ionospheric satellite transmitting harmonically related frequencies could be related to the electron content of the ionosphere up to the height of the satellite. The final equations contained approximate corrections to the first-order theory for various deviations from the elementary assumptions of the first-order theory.

This paper will present and discuss the results obtained by applying the method to the data recorded from satellite 1958 $\delta_2$  over a period of about 15 months extending from September 1958 through December 1959. The analysis has been limited to passes that occurred at or near local night hours during the early part of the period, because of difficulties arising from the keyed nature of the transmitter signals and because the satellite transmitter became inactive when the satellite was in the earth's shadow. Further limitations, by the considerations of part 1, are made to passes having heights in excess of 700 km and minimum zenith angles at the observing station less than 45°.

The passes to be considered here are shown

in Figure 1 as points whose position in local standard time and date are indicated on the coordinate axes. Each series of passes results from the westward precession of the orbital node in solar time, and, since the apsidal precession is small and the major axis only slowly varying, the passes comprising a series occur at approximately the same altitude, which is indicated on the figure.

Most of the observations were made at University Park, Pennsylvania (40.8°N 77.9°W), but the figure also includes some taken at the Ohio State University Radio Observatory (40.0°N 83.0°W) and some taken at the University of Virginia (38.0°N, 78.2°W).

The interruption in data from December 1958 to February 1959 was due to the poor quality of the satellite transmitter in this interval. The keyed pulses were typically of 100-millisecond duration or less and of irregular amplitude, making reliable scaling almost impossible.

**2. Experimental procedure.** The two satellite frequencies were received on narrow-bandwidth crystal-controlled receivers, and the audio beat notes obtained by injecting fixed, harmonically related, locally generated, stable reference frequencies into the inputs of the receivers were recorded on a multichannel tape recorder along with time information. To reduce these tape records to displays of dispersive Doppler frequency the signals are played through harmonically related automatic tracking heterodyne

<sup>1</sup>This paper was presented at the Symposium on the Exosphere and Upper *F* Region sponsored by Commission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.

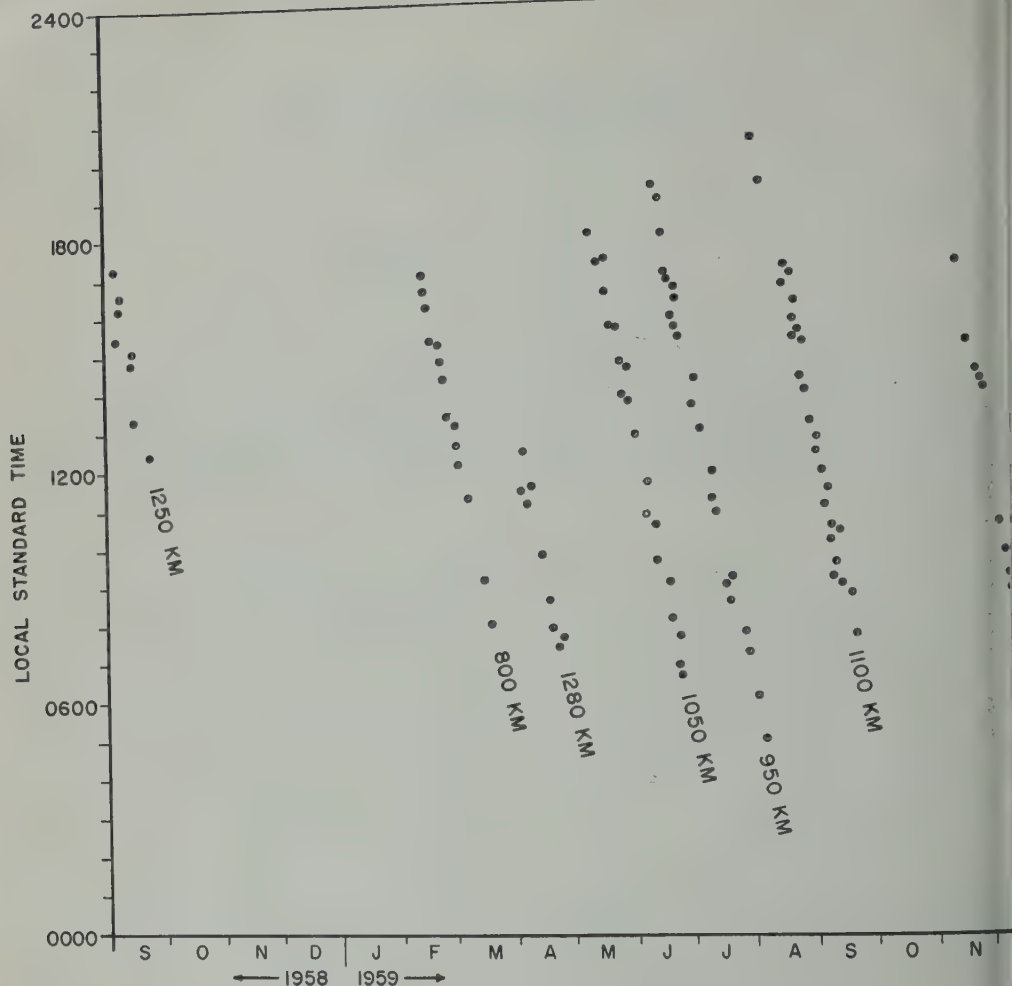


Fig. 1. Pass times of satellite 1958 $\delta_2$  included in this study.

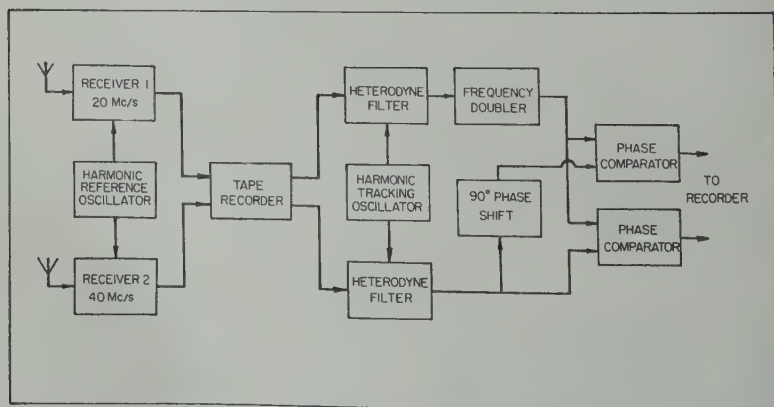


Fig. 2. Schematic diagram of the recording and analysis system for satellite 1958 $\delta_2$ .



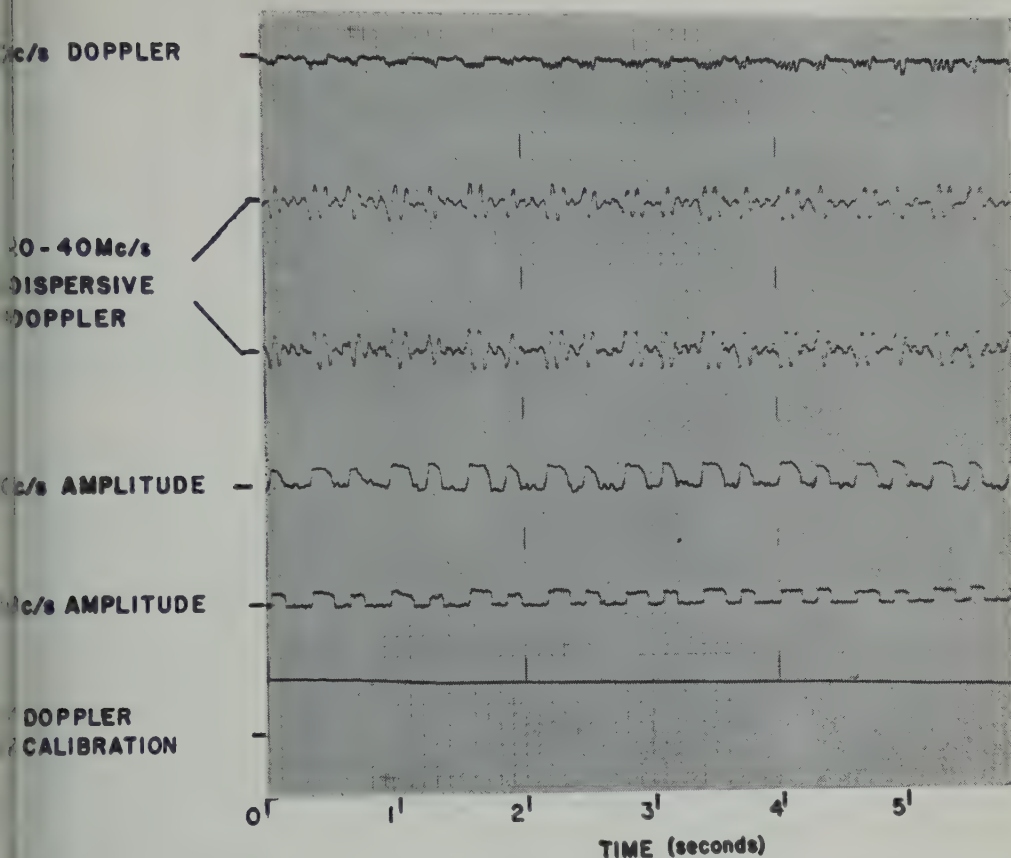


Fig. 3. Sample dispersive Doppler record of the 20- and 40-Mc/s signals from satellite 1958 $\delta_2$ .

whose bandwidth is about 100 c/s. The output from the 20-Mc/s channel is frequency-doubled and fed into quadrature phase comparators with the filter output from the 40-Mc/s channel, and the phase comparator outputs are then recorded on a multichannel magnetic recorder together with amplitude data from each channel, time information, and dispersive frequency data. The analysis system is shown schematically in Figure 2.

A section of sample record is shown in Figure 3. The keyed character of the satellite signals is clearly evident, and the quadrature dispersive Doppler channels contain the frequency information in keyed form also. Analysis of such data is possible even though the transmitter carrier is incoherent between pulses and 'pulls' the frequency during each pulse, because the dispersive Doppler frequency data are a plot of the frequency

$$(\dot{P}_2 - \dot{P}_1)/\lambda_{40}$$

which is a smoothly varying function.

Interpolation of the interpulse sections of dispersive Doppler data enables readings of dispersive Doppler frequency to be made throughout the record, which may last several hundred seconds. Frequencies can normally be read to an accuracy of  $\pm 0.1$  c/s average over about 1 second. They are plotted against time in the form shown in Figure 4; for comparison, the measured frequency at 20 Mc/s with an arbitrary zero of frequency is shown also.

The slope of the dispersive Doppler frequency curve through the zero of frequency is measured to give  $(\dot{P}_2 - \dot{P}_1)$ . The value of  $\dot{P}_1$  could be read from the slope of the 20-Mc/s Doppler frequency curve, but it has been found that frequency drifts are often present in the satellite transmitter and that this procedure may be

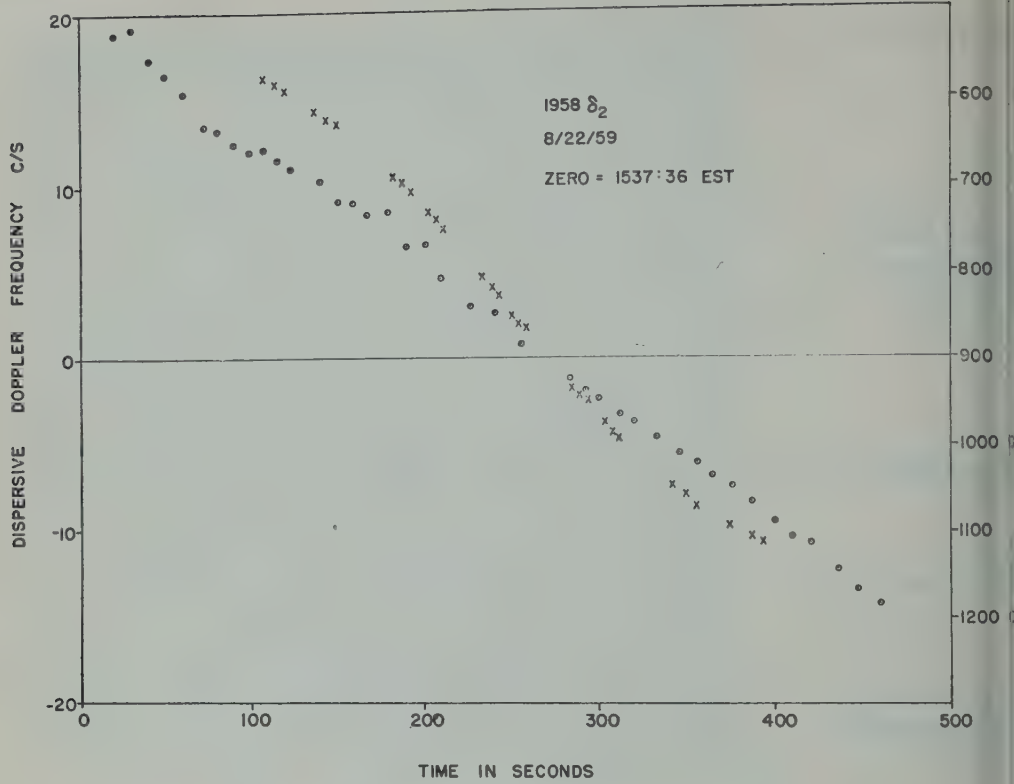


Fig. 4. Dispersive Doppler frequency (circles) and Doppler frequency (crosses) plotted for a typical satellite pass.

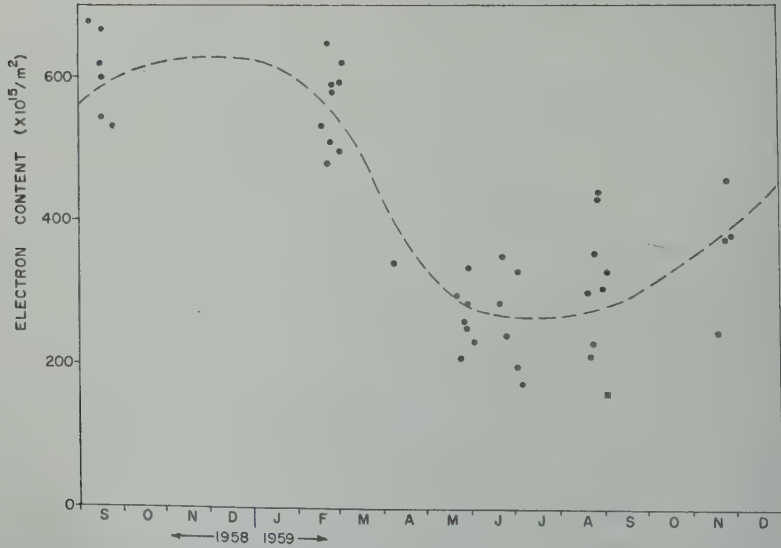


Fig. 5. Ionospheric electron content for satellite passes occurring in the period 1200-1600 LST.

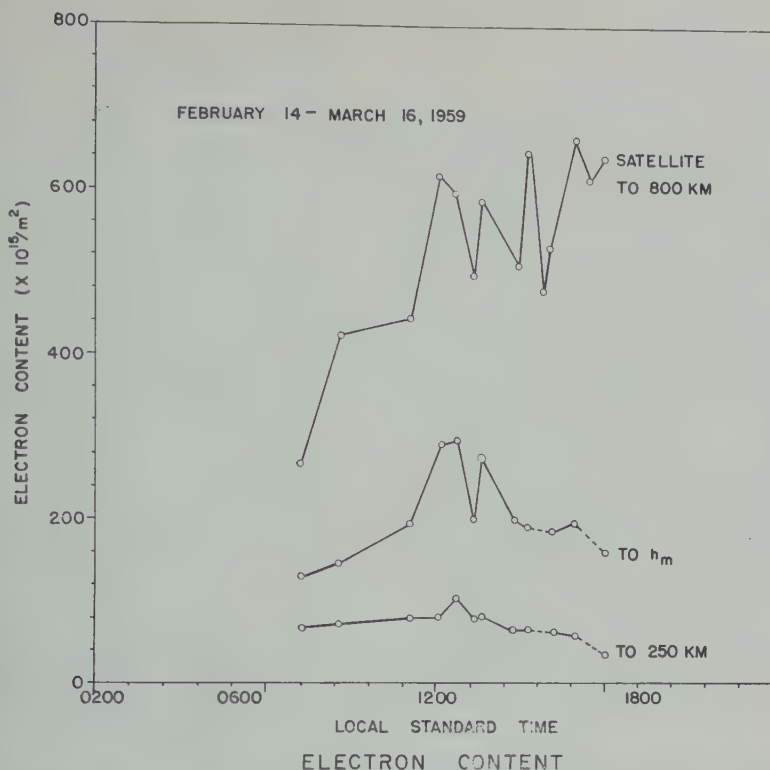


Fig. 4. Diurnal variation of electron content derived from the February-March 1959 series of satellite passes; mean altitude 800 km.

table. The value of  $\bar{P}_1$  is determined instead by computing the free-space rate of change of frequency,  $\bar{P}$ , from orbital data, and the value of  $\bar{P}_1$  can then be obtained to a good approximation from the equation

$$\bar{P}_1 = \bar{P} - 4(\bar{P}_2 - \bar{P}_1)/3$$

second term of which is measured as above. Once the value of  $R$  can be determined and combined with orbital data and ionosonde data by the methods developed in part 1 to give estimates of ionospheric electron content to the position of the satellite.

**Seasonal effects.** Figure 5 shows a plot of ionospheric electron contents calculated by the above methods for early afternoon passes (1200-1600 hours) when the ionospheric content is near its maximum for the particular day. A strong seasonal variation is evident with the same seasonal variation being seen near the *F*-layer peak. The variation in content from the winter to summer extremes is about 2.5:1. This result is the op-

posite of that found by *Evans* [1957] by lunar-radar techniques. It would be surprising to find any greatly different result, however, since the major contribution to the ionospheric electron content comes from the height range near the level of maximum density.

The only directly comparable results of systematically measured electron contents are those of *Garriott* [1960] using polarization rotation data from the same satellite considered here. His data overlap the period of this study but are centered about the winter solstice of 1958, and no strong seasonal variation is evident. *Garriott's* values of ionospheric electron content are appreciably greater than those in Figure 5, by about 15 to 20 per cent on the average, for the same satellite passes. This discrepancy may be due in part to the 3° lower geographic latitude and the 10° lower magnetic latitude of Stanford compared with University Park.

Measures of ionospheric electron content have been made by the lunar-radar technique by sev-



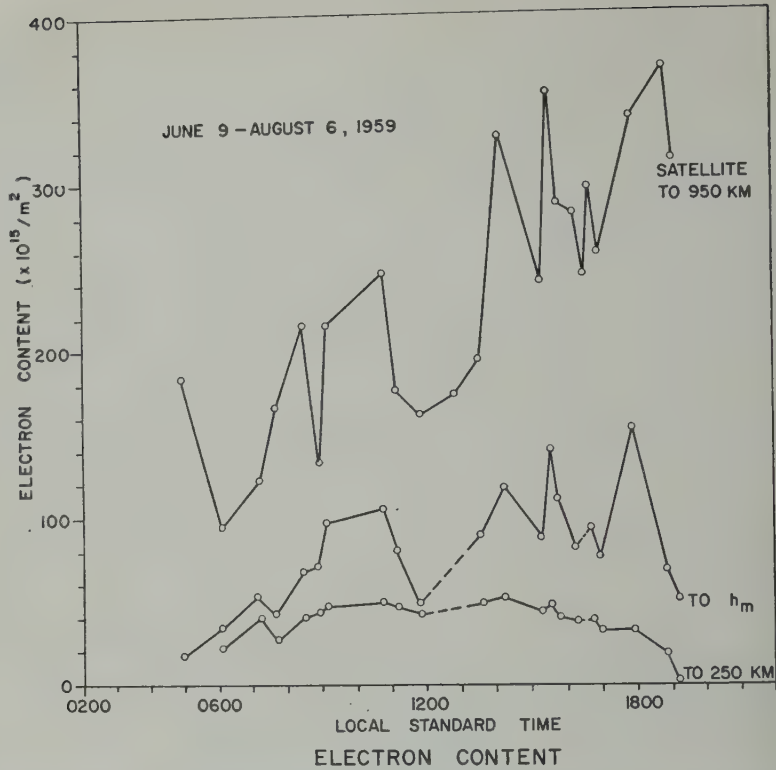


Fig. 7. Diurnal variation of electron content derived from the June-August 1959 series of satellite passes; mean altitude 950 km.

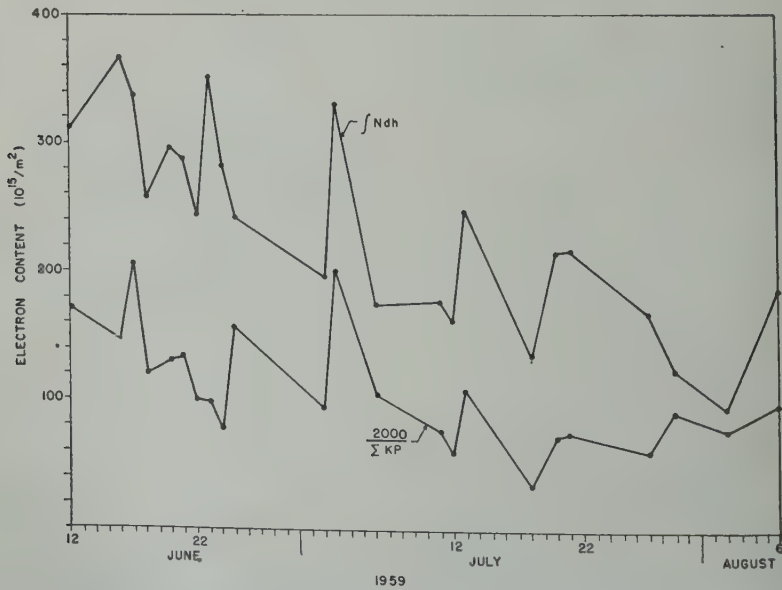


Fig. 8. Electron contents for June-August 1959 compared with the reciprocal of the sum of the primary 3-hour magnetic index,  $K_p$ , for the previous 24 hours.

workers [Evans 1957; Bauer and Daniels, 1959; Millman, Sanders, and Mather, 1960] but direct numerical comparison of their results with those of this paper is probably not possible either because of the greatly different geographic latitudes involved or because the measurements were restricted to nighttime only, which have not been evaluated here.

The apparent seasonal trend in Figure 5 may be entirely correct, as the content may also be steadily decreasing with the diminishing ionospheric activity through the period.

*Diurnal and day-to-day variations.* The electron content from two selected series of satellite passes has been plotted in Figures 6 and 7 with the local solar time as abscissas, and for comparison the electron contents, as determined from reduced concurrent ionograms, are also shown. Each point represents the information from one day's record.

General features of these plots are immediately apparent. First, there seems to be a steady average increase in total electron content from about sunrise through the day to near sunset although the lower-ionosphere data show the familiar late afternoon decrease in content. The implication is that the upper ionospheric electron content does not reach quasi-equilibrium until late afternoon after the lower ionosphere has already begun its evening decrease in content.

Second, there are day-to-day variations in electron content, which may be as large as  $\pm 20\%$  about the mean. That these variations are real and not due to random experimental errors may be seen from the fidelity with which they are reproduced in the lower-ionospheric data obtained from completely independent measurements, especially for the summer series in Figure 7. This result is again at variance with that of Evans [1957], who, however, ascribes the day-to-day constancy of the total ionospheric content as a means of resolving his experimental uncertainty and by this procedure may have reduced the larger day-to-day variations in his final

treated in more detail in a later paper. A tentative result has been obtained that the electron content varies in an inverse manner with the planetary magnetic disturbance index,  $K_p$ . Figure 8 shows the plot of electron contents from a summer series of satellite passes with the reciprocal of  $K_p$  summed over the preceding 24 hours. The similarity of the two sets of data is evident. This procedure has been found not to give such close correlation for the available winter results and must be regarded as still tentative. Garriott [1960] found no systematic dependence of electron content on magnetic index for the winter months of 1958-1959, which is not in conflict with our findings.

In addition to the close similarity between variations in the magnetic index and the electron content in Figure 8, there is also a steadily decreasing trend in both sets of data. This trend in content could be considered to be due to the normal diurnal variation when plotted as in Figure 7, but in view of the apparent dependence on magnetic index shown in Figure 8 this conclusion must be regarded with caution.

It would seem, from the apparent variation of the total ionosphere with season, with time of day, and with at least some effects of solar origin, that attempts to synthesize ionospheric density profiles from the electron content data based on observations spaced over a considerable period of time must be entered into with caution, and if possible large bodies of data from magnetically quiet days should be averaged to reduce the sampling errors in the method.

*5. Ionospheric profiles.* With the data now at hand, i.e., electron density measurements for the lower ionosphere from sounding methods and electron content measurements of the total ionosphere from satellite observations, it is possible to fit one parameter to any assumed model of the upper ionosphere for each satellite pass. Rather than select such a model, however, we shall draw some general conclusions that will hold for any later model that is chosen.

The simplest basic property that can be evaluated is the equivalent slab thickness of the ionosphere at the density corresponding to the  $F_2$ -layer maximum. This has been done for all the satellite passes occurring in the time interval 1200-1600 EST, and the results are shown in Figure 9. The equivalent slab thickness is seen to

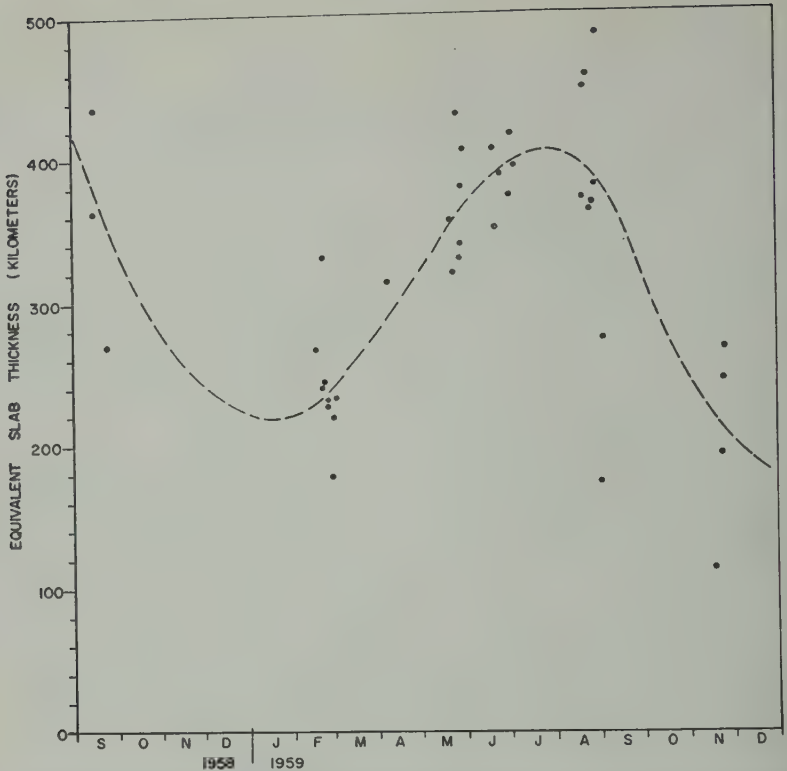


Fig. 9. Equivalent slab thickness of the total ionosphere for the period 1200-1600 EST.

range from a winter value of about 220 km to a summer value of about 400 km.

The constancy of the summer slab thicknesses for days on which the maximum density,  $N_o$ , varies by as much as 50 per cent indicates that in summer at least the ionospheric profile as a whole is not greatly distorted on magnetically disturbed days. No greatly depressed or enhanced maximum density values occurred on the winter days for which data are available, and the result of *Garriott* [1960] that the profile of the upper  $F$  region on magnetically disturbed days is appreciably broader than on quiet days in winter cannot be tested.

A more detailed analysis of the seasonal, diurnal, and day-to-day variations in electron content and electron density profile, their correlation with solar and geophysical data, and their interpretation in terms of ionospheric processes will form the basis of a later paper.

6. *Summary and conclusions.* The dispersive Doppler technique may be used with harmoni-

cally related satellite transmitters to evaluate electron content of the ionosphere, provided that suitable allowances are made for deviations of the actual ionosphere from the ideal model. The results obtained seem to be in close agreement with those obtained from independent polarization studies.

The seasonal variation of the total ionosphere has been measured; even though the summer-winter mid-latitude anomaly is only about that seen in the maximum electron density, the cause of seasonal changes in layer profile is still very marked. There remains the possibility that the magnitude of the seasonal variation is affected by the decreasing solar activity throughout the period of recording, and continued observations must be made to resolve this question.

The constancy of the ionospheric electron content from day to day has been demonstrated, although the gross shape of the layer has been shown to be almost constant, at least in summer.



ne. No direct inferences can be drawn of the detailed shape of the profile, however. It is questionable whether much significance can be attached to the commonly evaluated profile of electron contents above and below the maximum. The preliminary summary results given here indicate that, although the profile of the lower ionosphere varies considerably from day to day, the equivalent slab thickness of the total ionosphere does not, and therefore that many of the subpeak variations may be regarded as arising from relatively minor density deformations near the maximum density

The differences in the relationship of the ionosphere to magnetic index given by Garriott for winter and by this paper for summer support the conclusion of Maeda and Sato that the ionospheric response to magnetic activity varies with season in middle latitudes. Several large areas of study of the satellite ionosphere remain to be completed. The analysis of the same records from satellite 1958 $\delta_2$  has been very difficult because of the large scatter in the dispersive Doppler data. The use of continuously radiating satellite should reduce this difficulty. The analysis of records for satellite passes lying well in the ionosphere will be undertaken separately, as the validity of the methods used here has not yet been established under these conditions.

*Acknowledgments.* The cooperation of Dr. J. D. Kraus and the staff of the Ohio State University Radio Observatory and of Dr. E. C. Stevenson of the University of Virginia in providing facilities for satellite observations is greatly appreciated. The assistance of the technical and computing staffs of the Ionosphere Research Laboratory in recording and reducing the data is also acknowledged. Orbital data were provided by the Smithsonian Institute Astrophysical Observatory. Financial support was obtained from the National Science Foundation, the National Aeronautics and Space Administration, and the Wright Air Development Division under contract AF 33(616)-6157.

#### REFERENCES

- Bauer, S. J., and F. B. Daniels, Ionospheric parameters deduced from the Faraday rotation of lunar radio reflections, *J. Geophys. Research*, **63**, 439-442, 1958.
- Bauer, S. J., and F. B. Daniels, Measurements of ionospheric electron content by the lunar radio technique, *J. Geophys. Research*, **64**, 1371-1376, 1958.
- Evans, J. V., The electron content of the ionosphere, *J. Atmospheric and Terrest. Phys.* **11**, 259-271, 1957.
- Garriott, O. K., The determination of ionospheric electron content and distribution from satellite observations, 2, Results of the analysis, *J. Geophys. Research*, **65**, 1151-1157, 1960.
- Maeda, K. I., and T. Sato, The  $F$  region during magnetic storms, *Proc. IRE*, **47**, 232-239, 1959.
- Millman, G. H., A. E. Sanders, and R. A. Mather, Radar-lunar investigations at a low geomagnetic latitude, Paper presented at URSI meeting in Washington, D. C., May 4, 1960, *J. Geophys. Research*, **65**, 2619-2626, 1960.



## Faraday Rotation Observations of the Electron Content of the Exosphere<sup>1</sup>

R. B. DYCE

*Stanford Research Institute  
Menlo Park, California*

**Abstract.** The polarization twist imposed on 106.1-Mc/s radio waves by the ionosphere has been investigated by using the moon as a passive reflector. This has been done with a view to determining the total electron column density even at altitudes above the known ionosphere. Because the antenna is capable of being continuously directed at the moon for 12 consecutive hours, observations are possible from the predawn ionization minimum to the noontime maximum.

Due to the electron content above the peak of the  $F$  region is provided by a method that employs the measurement of the twist of polarization imposed by the ionosphere on a plane-polarized VHF wave reflected from the moon. The presence of the earth's magnetic field in the  $F$  ionosphere causes roughly ten complete rotations of the plane of polarization of a 100-Mc/s wave passing one way through the ionosphere. This gives the integrated electron content in the region where the earth's magnetic field is appreciable. This region can be considered to extend out to, say, 1 earth's radius above the earth's surface, for at that distance the earth's magnetic field has fallen to one-eighth its surface value. Originally it was assumed, however, it is known, that the electron density at the top of the  $F$  region diminishes to small values in a monotonic fashion. Thus it is relatively safe to consider the magnetic field constant through this region and to claim that we are measuring the electron content of the  $F$  region (the  $E$  region being negligible by comparison). Originally it was assumed that the magnetic field beyond the earth's field in interplanetary space was very small compared with its value at the earth's surface. This assumption has now been proved correct, so that the amount of

Faraday rotation introduced in interplanetary (cislunar) space appears to be small compared with that due to the  $F$  region, even though the distance to the moon is very large. When a linearly polarized radio wave strikes the moon's surface, a linearly polarized wave is returned. It can readily be shown that the cross polarization introduced over the entire circuit is of the order of 10 or 20 db below the primary polarization. Thus, the moon does not alter the polarization, although drastic changes may occur in the amplitude or in the position on the moon's surface from which reflection is obtained.

The earliest work was done at Jodrell Bank Experimental Station in England [Evans, 1957]. A single linearly polarized antenna was employed for transmission and reception for approximately 30 minutes each day. During a typical operating period, the received echo regularly changed amplitude several times. This was interpreted as a change in the amount of polarization twist. Two frequencies were used (100 and 120 Mc/s), and the sinusoidal oscillations of amplitude due to polarization rotation were observed to change phase with respect to each other. In this way, the ambiguity could be resolved with an uncertainty of only about one half-rotation. This equipment primarily suffered from its inability to track the moon over long periods of time each day.

Another successful experiment was accomplished by the United States Signal Corps and the University of Illinois on 151 Mc/s [Bauer and Daniels, 1959]. A steerable 50-foot dish was

<sup>1</sup>This paper was presented at the Symposium on the Exosphere and Upper  $F$  Region sponsored by Commission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.



used for transmitting (linearly polarized) and a steerable 28-foot dish (equipped with a slowly rotating polarization receptor) was used for receiving at Urbana, Illinois, several hundred miles away. This equipment could track the moon from moonrise to moonset on a given day, although observation during daytime hours was difficult on account of high electrical noise levels. Originally, ambiguity was resolved by considering the rate of change of electron density below the  $F$ -layer maximum observed by an ionospheric sounder. During later tests, a second frequency at 413 Mc/s was used over the same path, utilizing the fact that the Faraday rotation varies inversely as the square of the frequency.

Stanford Research Institute has also performed observations of the Faraday rotation on moon echoes, using a 60-foot steerable antenna on 106.1 Mc/s for both transmitting and receiving [Hill and Dyce, 1960]. Again, the moon could be tracked from moonrise to moonset. Only one frequency of operation was available. The antenna feed consisted of two dipoles mounted at right angles to each other, so that the observation of a minimum on one polarization channel could be correlated with the observation of a maximum on the other polarization channel. Measurements at any time of the day were possible from moonrise to moonset, the most striking effects being achieved at sunrise when the moon was at upper transit.

The ratio  $N_a/N_b$  (the electron content above the  $F$ -layer maximum to the electron content below) is given in Table 1.

Wright [1960] has shown that the ratio of  $N_a/N_b$  for a Chapman distribution is equal to 2.15. He points out, however, that this ratio may be changed by varying the assumed scale height as a function of altitude. Using a scale height of 100 km in the region above the maximum of the  $F$  layer, and the electron content observed by

ionosphere sounders below, Wright is able to duplicate the values of  $N_a/N_b$  appearing on the first two lines of Table 1.

TABLE 1. Observed Values  $N_a/N_b$  by the Lunar-Echo Method

| Observer                 | Night  | D   |
|--------------------------|--------|-----|
| Evans [1957]             | 4      | ≤ 2 |
| Bauer and Daniels [1959] | 4 or 5 |     |
| Hill and Dyce [1960]     | 1.5    |     |

The moon-echo method of measuring the electron content of the exosphere was an early indication that the  $F$  layer did not have the shape of a symmetrical parabola, but that greater electron content existed above the ionization maximum than below, by about a factor of 3. This result has since been confirmed by other methods using satellites, rockets, and incoherent scatter radars. The advantage in the moon-echo method lies in the moderate power requirements and the fact that observations can be made a continuous 12 hours and from one day to the next.

It is unfortunate that none of the observations to date have captured the onset of a magnetic storm.

#### REFERENCES

- Bauer, S. J., and F. B. Daniels, Measurement of ionospheric electron content by the lunar echo technique, *J. Geophys. Research*, **64**, 1371-1375, 1959.
- Evans, J. V., The electron content of the ionosphere, *J. Atmospheric and Terrest. Physics*, **25**, 259-271, 1957.
- Hill, R. A., and R. B. Dyce, Some observations of ionospheric Faraday rotation on 106.1 Mc/s, *J. Geophys. Research*, **65**, 173-176, 1960.
- Wright, J. W., A model of the  $F$  region at high latitudes, *J. Geophys. Research*, **65**, 185-191, 1960.

# Radar-Lunar Investigations at a Low Geomagnetic Latitude<sup>1</sup>

GEORGE H. MILLMAN AND ARMAND E. SANDERS

*General Electric Company  
Syracuse, New York*

AND

ROBERT A. MATHER

*Rome Air Development Center  
Rome, New York*

**Abstract.** Radar reflections from the moon have been studied utilizing the Trinidad, BWI, radar operating in the 400-Mc/s frequency range. The moon has been tracked from moonrise through transit during periods of sunset. On all occasions, the rotation of the plane of polarization (Faraday effect) was observed. Absolute values of the total electron content of the ionosphere deduced from these measurements and the electron content above the height of the  $F$ -layer maximum, evaluated from vertical-incidence ionospheric soundings recorded at the NBS Puerto Rican station, are presented.

**Introduction.** When a linearly polarized electromagnetic wave is propagated through an anisotropic medium immersed in a magnetic field, the plane of polarization undergoes a rotation. This phenomenon, commonly referred to as the Faraday effect, is known to take place for radio waves traversing the ionosphere.

Radar techniques employed to reflect signals from the moon's surface in the study of the characteristics of the ionosphere by means of the Faraday effect were first reported by *Evans* [1956, 1957]. By comparing the signal-amplitude fading at two closely spaced frequencies, 120.0 Mc/s and 120.5 Mc/s, per cent, Evans was able to deduce the ionospheric electron content during the periods when the moon crossed the meridian at Manly, England.

*Bauer and Daniels* [1958, 1959; and *Daniels and Bauer*, 1958] performed a similar measurement at a frequency of 151.11 Mc/s, employing a bistatic radar system operating between Belvidere, New Jersey, and Urbana, Illinois. The absolute values of the total electron content were

evaluated from these measurements by utilizing ionospheric electron-density profiles below the  $F$ -layer maximum derived from vertical-incidence sounding data.

The recent work of *Hill and Dyce* [1960] at Menlo Park, California, on 106.1 Mc/s was complicated by inability to determine the absolute electron content because of the ambiguity in resolving the number of polarization rotations of the wave.

The investigation reported in this paper was conducted at Trinidad, BWI (22° geomagnetic north latitude), with a pulsed radar operating at a frequency of approximately 400 Mc/s. Although only one frequency was employed, the absolute number of rotations, and thus the total ionospheric electron content, was resolvable.

The moon was tracked from moonrise through transit on January 12 and February 8, 1960. These dates correspond to the time when the moon was at its maximum northern declination.

A selected sample of an amplitude-versus-time film record of moon echoes recorded at Trinidad and displaying Faraday rotation is illustrated in Figure 1. It is seen that the reflected signal is predominantly oriented in the vertical-polarization receiver channel during the period when horizontal polarization is employed on transmission.

<sup>1</sup>This paper was presented at the Symposium on the Ionosphere and Upper  $F$  Region sponsored by Mission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.

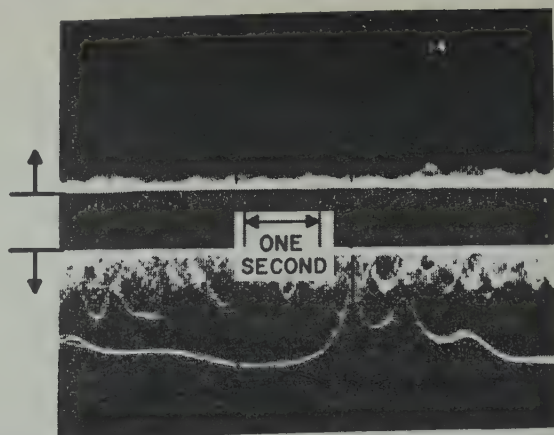
HORIZONTAL  
POLARIZATIONLOGARITHM  
OF AMPLITUDEVERTICAL  
POLARIZATIONLOGARITHM  
OF AMPLITUDE

Fig. 1. Selected sample of amplitude vs. time film record of moon echoes taken at Trinidad displaying Faraday rotation; February 8, 1960, 1404 EST; 400 Mc/s frequency range; horizontal polarization transmitted.

*Mathematical considerations.* The polarization rotation experienced by a linearly polarized wave traversing a two-way path in the ionosphere is represented by the function

$$\phi = \frac{4.7233 \times 10^4}{f_0^2} \int_{h_1}^{h_m} f(h) H \cos \theta N dh \quad (1)$$

where  $\phi$  is the rotation in radians,  $f_0$  is the transmission frequency in cycles per second,  $H$  is the magnetic field in gauss,  $N$  is the electron density in electrons per cubic centimeter,  $dh$  is the height differential in centimeters, and  $\theta$  is the propagation angle, i.e., the angle between the direction of the earth's magnetic lines of force and the direction of electromagnetic propagation.

The function  $f(h)$ , which is basically the secant of the angle between the ray path and the zenith, is given by

$$f(h) = \frac{r_0 + h}{[(r_0 + h)^2 - (r_0 \cos E)^2]^{1/2}} \quad (2)$$

where  $r_0$  is the radius of the earth and  $E$  is elevation angle of the antenna beam.

Equation 1 is representative of the quasi-longitudinal mode of propagation. For frequencies in the 400-Mc/s range, this relationship is valid for  $0 \leq \theta \leq 89^\circ$ .

It should be noted that the magnitude of the Faraday rotation,  $\phi$ , is ambiguous in that

$$\phi = n\pi \pm \Delta\phi \quad (3)$$

where  $n$  is a positive integer and  $\Delta\phi$  is the acute

polarization rotation angle normally indicated in a radar-lunar measurement.

For transmission to the moon, equation 1 must be modified to the form

$$\phi = \frac{4.7233 \times 10^4}{f_0^2} \left[ \int_0^{h_m} f(h) H \cos \theta N dh + \overline{[f(h) H \cos \theta]}_{h_m}^\infty \int_{h_m}^\infty N dh \right]$$

where  $h_m$  is the height of the  $F$ -layer maximum and

$$\overline{[f(h) H \cos \theta]}_{h_m}^\infty$$

is the average value of the function  $f(h)H \cos \theta$  between the limits of the height of the  $F$ -layer maximum and the distance to the moon. As a first approximation, it is valid to assume that

$$\overline{[f(h) H \cos \theta]}_{h_m}^\infty \approx \overline{[f(h) H \cos \theta]}_{h_m}^{1000 \text{ km}}$$

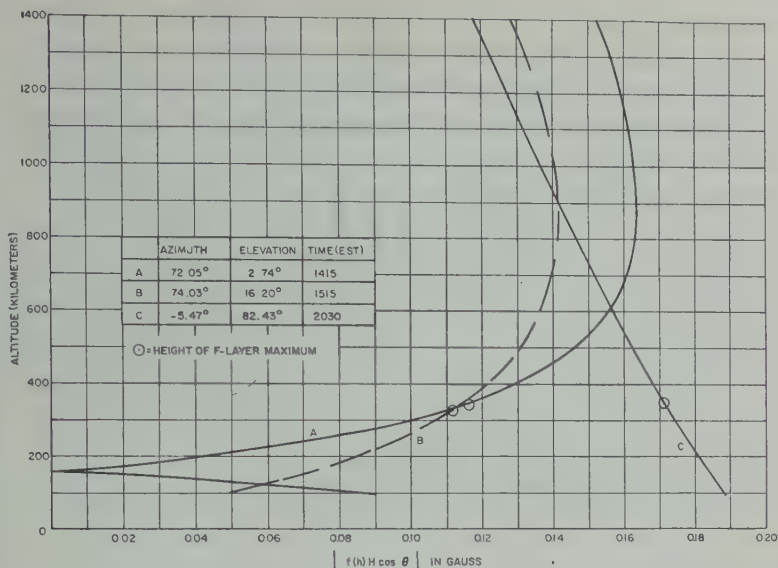
since the contribution of the integrated electron density to the angular rotation of the wave is considered to be negligible above a height of 1000 km [Evans, 1956].

The electron content above the  $F$ -layer maximum can be expressed in terms of the electron content below the  $F$ -layer maximum by the constant  $K$ , where

$$\int_{h_m}^\infty N dh = K \int_0^{h_m} N dh$$

The absolute magnitude of  $f(h) H \cos \theta$



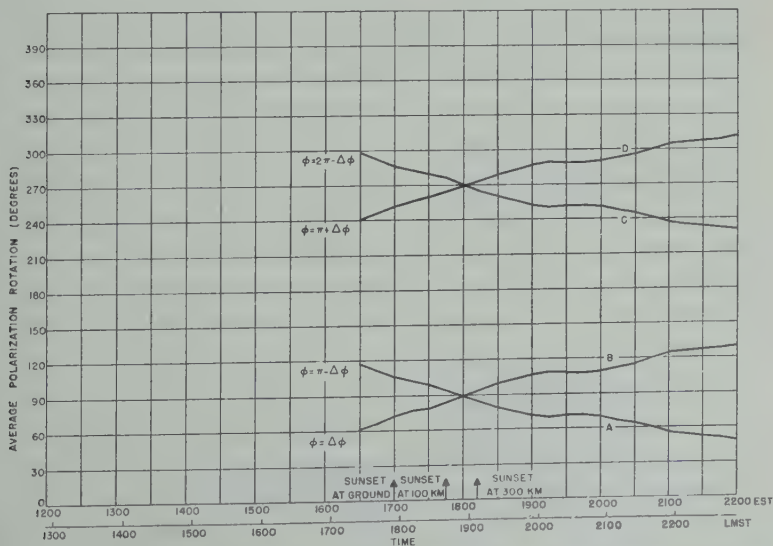


The function of  $f(h) H \cos \theta$ , for the moon orbit of February 8, 1960, as viewed from Trinidad.

from Trinidad is plotted as a function of  $f(h) H \cos \theta$  in Figure 2 for the lunar orbit of February 8, 1960. These calculations are based on the assumption that the earth's magnetic field can be approximated by a centered magnetic field having a geomagnetic pole at 78.6°N and 10.1°W. The method for determining the

propagation angle,  $\theta$ , at ionospheric heights has been discussed by Millman [1959].

It is evident that, at low angles of elevation, the function can vary over quite a wide range. As the elevation is increased, there is a noticeable reduction in the spread of its value as a function of height.



3. Average Faraday rotation from Trinidad radar-lunar observations, January 12, 1950.

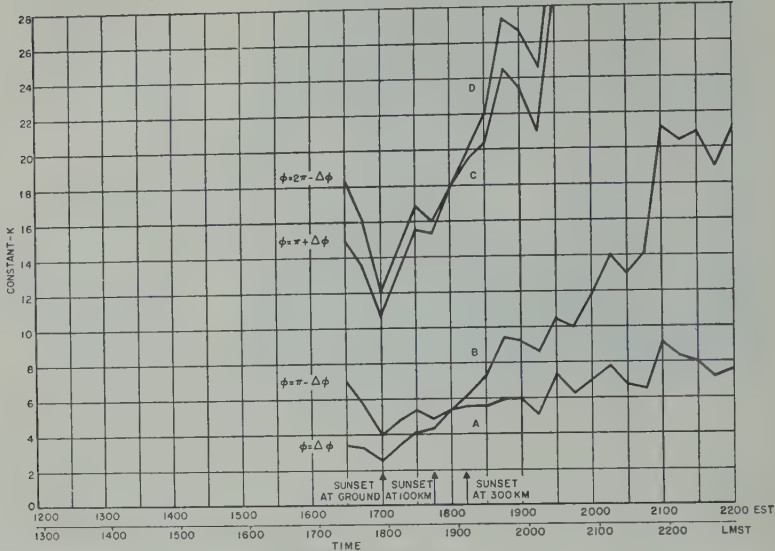


Fig. 4. Variation of the ratio of the electron content above the *F*-layer maximum to the electron content below the *F*-layer maximum, January 12, 1960.

Because of the extreme variation in magnitude at low elevation angles, it is necessary to retain  $f(h) H \cos \theta$  within the integral up to the *F*-layer maximum height. Above this region, the function can be approximated by an average value taken between the *F*-layer maximum and the 1000-km height limits (see Fig. 2).

*Analysis of experimental data.* The average

acute angular rotation measured during radar-lunar observation of January 12, 1960 is presented by the  $\phi = \Delta\phi$  and the *A* curves in Figure 3. Because of the inability of an antenna receiver system to distinguish the true orientation, it is necessary to consider the other possible values of the polarization angle given by equation 3 and plotted in Figure 3.

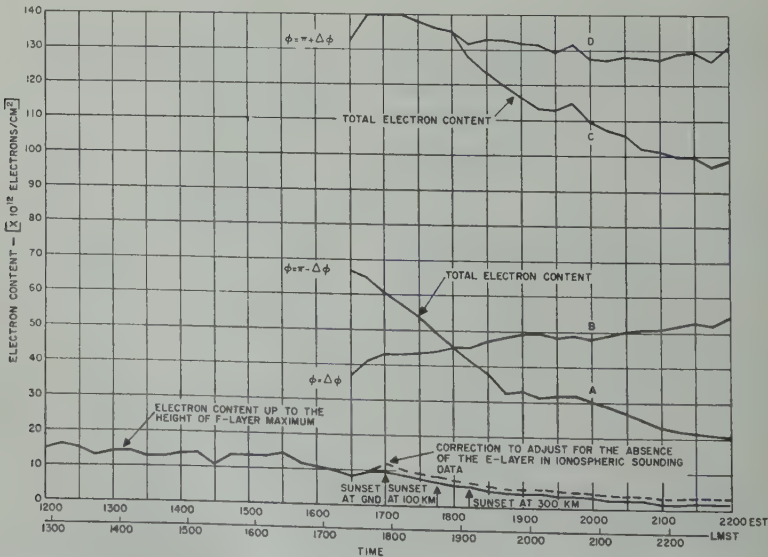


Fig. 5. Electron content in a 1 sq cm vertical column through the ionosphere, January 12, 1960.

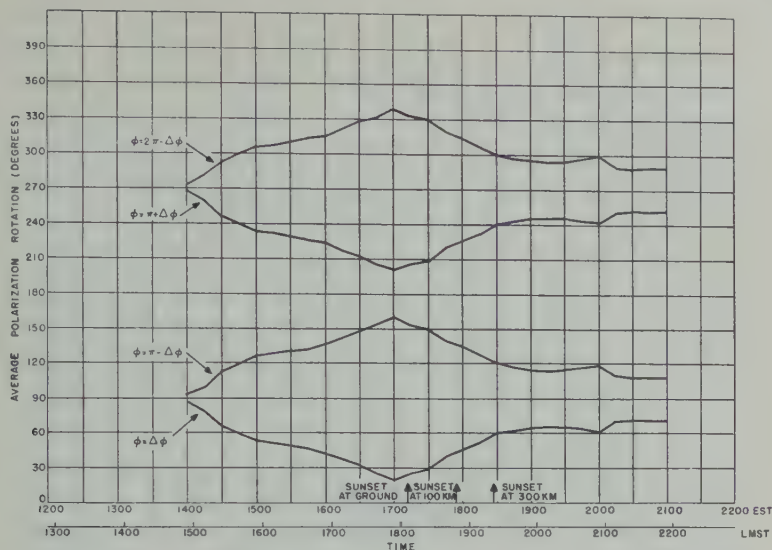


Fig. 6. Average Faraday rotation from Trinidad radar-lunar observation, February 8, 1960.

the intersection of the  $\phi = \pi + \Delta\phi$  and  $\phi = \pi - \Delta\phi$  curves, as at 1800 EST, also occurs when the polarization of the received signal passes through either a  $90^\circ$  or a  $180^\circ$  angular rotation. It should be noted that an intersection could result either from a phase reversal or from a continuing monotonic phase change. For example, there are two possible paths which each of the  $\phi$  curves could follow; the  $\phi = \Delta\phi$  curve could have continued along the B path instead of the A path, and continuity of phase suggests that this is indeed the case.

Figure 4 is a plot of the ratio of the integrated electron content above the *F*-layer maximum to the integrated electron content below the *F*-layer maximum. It is of interest to note that  $K$  varies as a function of time, its magnitude being larger at sunset than before. The vertical-incidence ionospheric sweep-frequency data recorded at the National Bureau of Standards, Puerto Rico Station, were utilized in the calculations of this parameter. The conversion of the virtual height versus frequency records into true height versus electron-density profiles was performed by means of the Budden [1955] matrix technique, the components of the matrix being furnished by E. R. Merling of the Pennsylvania State University. At this point in the analysis, it is still not possible to deduce the correct time variation of  $K$

throughout the observational period, but it appears quite improbable, on the basis of lunar results of Bauer and Daniels [1959] and the analysis of satellite transmissions by Garriott [1960a, b], that  $K$  could attain the extremely high values indicated by the *C* and *D* curves.

The electron content in a 1 sq cm vertical column through the ionosphere is depicted in Figure 5. The *B* curve would indicate that the total ionospheric electron content was increasing during the time the electron content below the height of the *F*-layer maximum was decreasing. This would require that the electron production above the *F*-layer maximum be continued throughout the night at a greater rate than the electron loss below the *F*-layer maximum. Since the mechanism governing this physical process is difficult to visualize, it is believed permissible to discard the *B* data entirely.

According to Figure 5, therefore, the total ionospheric electron content is definable by either the  $\phi = \Delta\phi$  or the  $\phi = \pi - \Delta\phi$  curve together with the *A* curve. A comparison of these results with the data presented below suggests that  $\phi = \pi - \Delta\phi$  is the more likely choice.

The average Faraday rotation observed during the lunar orbit of February 8, 1960, is shown in Figure 6. For this particular example, the



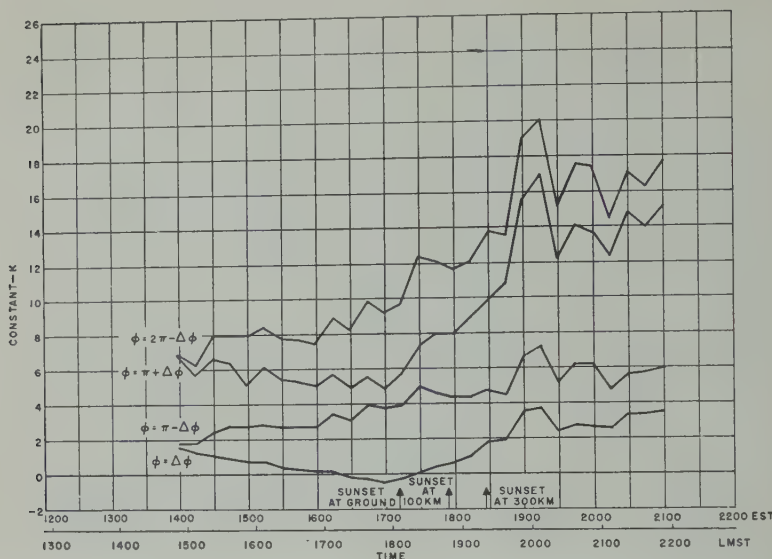


Fig. 7. Variation of the ratio of the electron content above the  $F$ -layer maximum to the electron content below the  $F$ -layer maximum, February 8, 1960.

acute polarization angle reaches a minimum of about  $20^\circ$  and then reverses its direction. Near moonrise, the signal undergoes approximately a  $90^\circ$  angular rotation even though the direction of propagation satisfies the requirement of perpendicularity with the magnetic field in the vicinity of the 160-km height level. This discrepancy merely indicates that quasi-transverse propagation is attained only at a small region along the transmission path and that the quasi-longitudinal mode of propagation is predominant everywhere else.

An analysis of the data shown in Figure 7 reveals that the  $\phi = \Delta\phi$  curve can be eliminated at once. This conclusion is arrived at simply because the ratio of the electron content above the  $F$ -layer maximum to the electron content below goes to zero and also becomes negative.

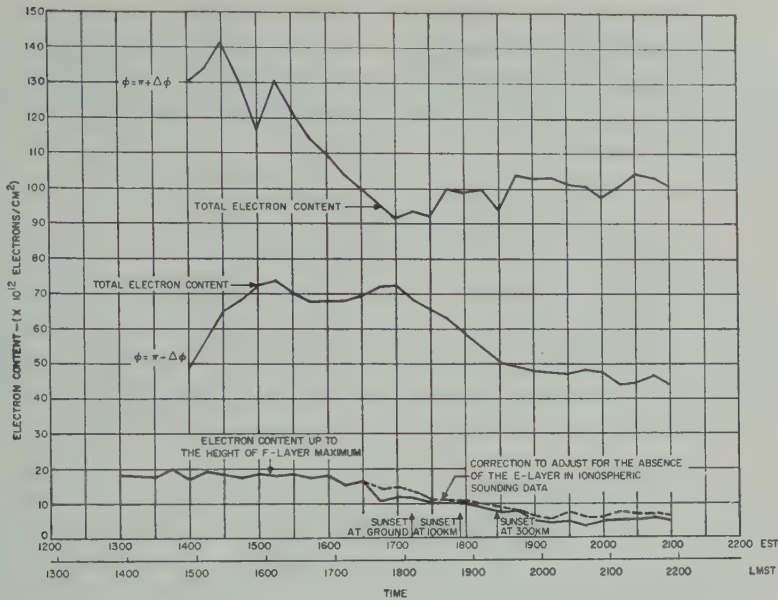
An examination of Figure 8 indicates that the  $\phi = \pi - \Delta\phi$  curve is the most likely correct solution to the problem. The general shape of this curve strongly resembles the  $\phi = \pi - \Delta\phi$  curve of Figure 5. The increase in total electron content between 1400 and 1515 EST implies that high electron production was taking place only in the region above the  $F$ -layer maximum while electron production below the peak of the  $F$ -layer remained comparatively constant.

The  $\phi = \pi + \Delta\phi$  curve of Figure 8 is not believed to be a reasonable solution, simply because of the tendency of the total electron content to increase while the electron content measured by ordinary ionosphere soundings tends to decrease toward the end of the interval. Both it and the  $\phi = 2\pi - \Delta\phi$  curve are already highly suspect because of the large values of  $K$  they provide in Figure 7.

**Conclusions.** It is shown that, by a process of elimination, it is possible to arrive at a reasonable solution to the polarization ambiguity problem usually present when only one frequency is employed in the study of radar-lunar reflectivity.

An interesting conclusion resulting from the experiment is that, for transmission through the entire ionosphere, the condition of complete transverse propagation is maintained in only a very small region in space. The Faraday rotation, even when the condition of perpendicularity of the propagation direction with the magnetic field is satisfied, can still be appreciable.

The ratio of the integrated electron density above the peak of the  $F$ -layer maximum to that below appears to be of the order of 3 to 5 during the daylight hours before sunset and decreases to approximately 6 to 8 in the early hours of the evening. It is estimated that



8. Electron content in a 1 sq cm vertical column through the ionosphere, February 8, 1960.

results could have an rms error between approximately 0.5 and 0.7 and the latter about 1.5.

The ratio of the electron contents as measured by Bauer and Daniels [1959] compares rather favorably with the data presented in this paper; [1957] results, for the nighttime conditions, are slightly lower.

The measurements of Hill and Dyce [1960], on the other hand, are considerably lower than the Trinidad data. This discrepancy is most likely brought about by their assumption that the function  $f(h) H \cos \theta$  can be considered a constant in evaluating the electron-content inte-

gral. From the preliminary data reported herein, it is obvious that the electron content above the F-layer maximum is a variable function of time. In order to study the characteristics of its diurnal variation, it is essential that radar-lunar observations be conducted in such a way that the electron content be tracked from moonrise through moon-

set. **Acknowledgments.** This work would not have been possible without the cooperation of innumerable General Electric and RADC personnel. The assistance of F. Rose in the data reduction is greatly appreciated. We wish to thank J. W.

Wright of the Central Radio Propagation Laboratory, National Bureau of Standards, for supplying the Puerto Rico ionosonde data.

This work is sponsored by the Rome Air Development Center under contract AF30(602)-1625.

#### REFERENCES

- Bauer, S. J., and F. B. Daniels, Ionospheric parameters deduced from the Faraday rotation of lunar radio reflections, *J. Geophysical Research*, **63**, 439-442, 1958.
- Bauer, S. J., and F. B. Daniels, The measurement of ionospheric electron content by the lunar radio technique, *J. Geophysical Research*, **64**, 1371-1376, 1959.
- Budden, K. A., A method for determining the variation of electron density with height ( $N(Z)$  curves) from curves of equivalent height against frequency, *The Physics of the Ionosphere*, pp. 332-339, The Physical Society, London, 1955.
- Daniels, F. B., and S. J. Bauer, Measurement of the ionospheric Faraday effect by radio waves reflected from the moon, *Nature*, **181**, 1392-1393, 1958.
- Evans, J. H., The measurement of the electron content of the ionosphere by the lunar radio echo method, *Proc. Phys. Soc. London, B*, **69**, 953-955, 1956.
- Evans, J. H., The electron content of the ionosphere, *J. Atmospheric and Terrest. Phys.*, **11**, 259-271, 1957.
- Garriott, O. K., The determination of ionospheric electron content and distribution from satellite

- observations, 1, Theory of the analysis, *J. Geophys. Research*, 65, 1139-1150, 1960a.
- Garriott, O. K., The determination of ionospheric electron content and distribution from satellite observations, 2, Results of the analysis, *J. Geophys. Research*, 65, 1151-1157, 1960b.
- Hill, R. A., and R. B. Dyce, Some observations of ionospheric Faraday rotation on 106.1 Mc, *Geophys. Research*, 65, 173-176, 1960.
- Millman, G. H., The geometry of the earth's magnetic field at ionospheric heights, *J. Geophys. Research*, 64, 717-726, 1959.



# Use of the Incoherent Scatter Technique to Obtain Ionospheric Temperatures<sup>1</sup>

T. E. VANZANDT AND K. L. BOWLES

National Bureau of Standards  
Boulder, Colorado

**Abstract.** If the ion-electron gas is in diffusive equilibrium on the topside of the  $F$  layer, then the electron density decreases exponentially with height, and its logarithmic decrement is proportional to the neutral gas temperature. From an electron density profile obtained by the scatter radar technique, it is shown that this interpretation is consistent. Moreover, the reduction of ionospheric temperatures in this way from scatter radar electron density profiles has several advantages over other methods.

The effectiveness of the incoherent scattering of radio waves by free electrons [Gordon, 1958] as a means of measuring the electron-density distribution with height above the earth's surface (called the 'electron-density profile') has been demonstrated by several experiments performed by the National Bureau of Standards [Vanzandt, 1958, 1960] and by Lincoln Laboratory [Pineo, Kraft, and Briscoe, 1960]. The present note reports on the physically interesting aspects of one set of the few profiles available at present.

Figure 1 the data from five soundings obtained between 1240 and 1315 local time on November 22, 1959, at the NBS station near Havana, Illinois, have been plotted on the same logarithmic scale. Each point shown in the figure represents the 5-minute average of incoherently scattered power from a height region 18 km thick, multiplied by the square of the range in order to correct for the variation of radar sensitivity with height. The lower solid curve is the electron-density profile deduced by the 'true-height' technique from the ionosonde record taken at 1300 LT at Havana, Illinois. The upper straight line is an example of a potential whose significance will be discussed later. Several features of this profile are worth noting.

This paper was presented at the Symposium on the Exosphere and Upper  $F$  Region sponsored by the Commission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.

First, although the scatter radar profile has been fitted to the ionosonde profile as well as possible, the fit is poor at the lower heights and again near  $h_{max}F_2$ . The departures at the lower end are probably due to echoes scattered coherently from each of many irregularities whose origin may be explained by turbulence, gravity waves, or similar phenomena. The low end may,

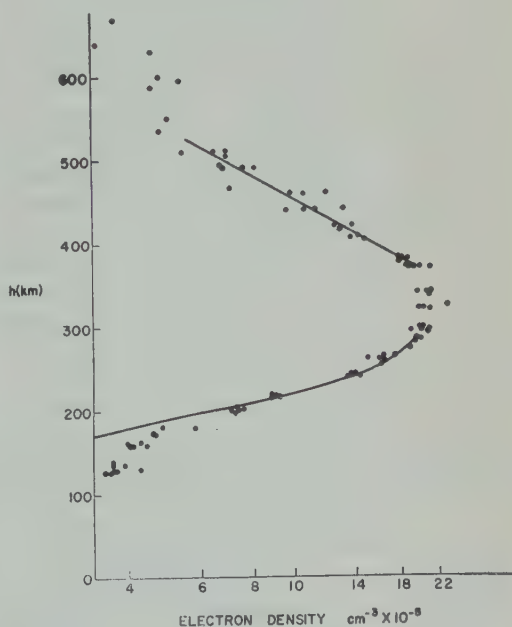


Fig. 1. The scatter radar data from 1240 to 1315 LT normalized to fit the electron-density profile (the lower solid curve) deduced from the ionogram for 1300 LT, Havana, Illinois, November 22, 1959.

in addition, indicate too much scatter echo due to side-lobe response of the antenna. However, it was shown by Bowles [1958, 1960] that at greater heights the scatter radar points are proportional to the electron density, and, indeed, from 200 to 280 km the scatter radar profile fits the ionosonde profile.

The departures near  $h_{\max} F_2$  may well be due to a small error in the height scale of the C-2 ionosonde or to the 'valley error' in the reduction of the ionogram to the electron-density profile, which can have a rather large effect on the goodness of fit.

The apparent bluntness of the nose of the  $F_2$  layer on the scatter radar profile may be a spurious effect caused by systematic differences among the electronic averaging devices. Such effects are evident even where the scatter radar data and the ionosonde profile fit well, between 200 and 280 km, since a smooth curve drawn through the scatter radar data would give rise to conspicuous ledges on the virtual height curve, which were not observed.

From the point of view of the physics of the upper atmosphere, the most interesting feature of the scatter radar profile is the exponential decrease of electron density between 370 and 520 km. (Whether the departure from this exponential above 520 km is real or is due to the limit of the experimental sensitivity cannot be determined from the present data.) Since the charged particles on the top side of the  $F$  region must be in diffusive equilibrium and since the atmosphere there is nearly isothermal the electron density decreases exponentially with a scale height twice that of the neutral atoms, which are probably atomic oxygen [Johnson, 1960]. Then the logarithmic decrement of  $128 \pm 11$  km deduced from the scatter radar profile leads to a temperature of  $1050^\circ \pm 90^\circ\text{K}$ . Such a value is within the experimental errors of  $F$ -region temperatures deduced from the widths of the 6300 Å nightglow [Mulyarchik, 1960; Wark, 1960] and from satellite drag data [e.g., Champion and Minzner, 1959]. Of course, since the temperature varies with height, time of day, latitude, and solar activity, such a comparison can be rough, at best.

This method of deducing the temperature from the logarithmic decrement of the electron density on the top side of the  $F$  layer should be applicable at all times of the day and night except near sunrise and sunset, and at all latitudes and high latitudes. At low latitudes the ionospheric gas is not in a vertical diffusive equilibrium. The present method has certain advantages over other methods. The Doppler width method requires long exposures at night and yields a temperature averaged over time and at an uncertain height in the  $F$  region. The satellite drag method requires weeks to months to go through one year of local time, and it consequently averages out variations of season and solar activity. In view of the advantages of the present method, it appears that the scatter radar technique may be invaluable for the determination of atmospheric temperatures.

*Acknowledgment.* We wish to thank J. Wright for his help in operating the C-2 ionosonde at Havana, Illinois, and for reducing the ionogram to electron-density profiles.

#### REFERENCES

- Bowles, K. L., Observation of vertical incoherent scatter from the ionosphere at 41 Mc/sec, *Rev. Letters*, **1**, 454-455, 1958.
- Bowles, K. L., Incoherent scattering by free electrons as a technique for studying the ionosphere and exosphere, *J. Research NBS*, part 2, in press, 1960.
- Champion, K. W., R. A. Minzner, Atmospheric densities from satellites and rocket observations, *Planetary and Space Sci.*, **1**, 259-264, 1958.
- Gordon, W. E., Incoherent scattering of radio waves by free electrons with application to space exploration by radar, *Proc. IRE*, **46**, 1829, 1958.
- Johnson, F. S., The ion distribution above the maximum, *J. Geophys. Research*, **65**, 571, 1960.
- Mulyarchik, T. M., Interferometric temperature measurements of the upper atmosphere based on the width of certain emission lines, *Proc. Sci., USSR*, Geophys. Ser., no. 3, pp. 44-45, 1960.
- Pineo, V. C., L. G. Kraft, and H. W. Briscoe, Spheric backscatter observation at 440 Mc, *J. Geophys. Research*, **65**, 1620-1621, 1960.
- Wark, D. Q., Doppler widths of the atomic oxygen lines in airglow, *Astrophys. J.*, **131**, 491, 1960.

## Some Characteristics of Ionospheric Backscatter Observed at 440 Mc/s<sup>1</sup>

V. C. PINEO, L. G. KRAFT, AND H. W. BRISCOE

*Lincoln Laboratory,<sup>2</sup>  
Massachusetts Institute of Technology  
Lexington 73, Massachusetts*

**Abstract.** Incoherent ionospherically backscattered radar returns have been observed at 440 Mc/s over an approximate height range of 100 to 800 km. The results of experiments designed to determine the variation with height of both the intensity and the frequency spectrum of backscattered returns are presented. Backscatter signal-intensity measurements at radio frequencies much greater than the  $F_2$  critical frequency provide an independent means of estimating electron densities. Height profiles of electron densities obtained from experimental data are discussed. Experimentally determined values of the scattering cross section are approximately equal to the square of the classical electron radius. A typical result of the frequency spectrum measurements at 440 Mc/s indicates a half-power spectrum width of about 11 kc/s at a height of about 300 km during midafternoon of a March day. Spectrum widths observed at several heights and at various times of day and night are discussed.

**Introduction.** Preliminary reports [Pieno, Kraft, and Briscoe, 1960] of an ionospheric backscatter experiment performed at 440 Mc/s at the MIT Lincoln Laboratory Millstone Hill Radar Station gave two principal results: (1) a Doppler  $F$ -region frequency spectrum about 10 kc/s wide at the half-power point, and (2) a height profile of backscattered signal power and a height profile of relative electron density.

The technique used in this experiment was similar to that suggested by Gordon [1958]. Gordon [1958] has reported results from a similar experiment at 41 Mc/s.

In this paper some new data are presented and the analysis of data previously shown is extended to the calculation of the cross section of the scatterers.

**Experimental details.** The location of the site and the operating parameters for the experiment are given below.

**Location:** Westford, Mass., 71.5°W, 42.6°N.  
**Frequency:** 440 Mc/s.

This paper was presented at the Symposium on the Exosphere and Upper  $F$  Region sponsored by Commission III of the United States National Committee of the International Scientific Radio Union (URSI), held on May 4, 1960, in Washington, D. C.

Supported with support from the U. S. Army, Navy, and Air Force.

**Antenna:** 84-ft-diameter paraboloid.  
**Antenna efficiency:** about 50 per cent.  
**Peak transmitted power:** 1.6 megawatts.  
**Pulse duration:** 500 microseconds.  
**Pulse-repetition frequency:** 50 pps.  
**Transmitted polarization:** right circular.  
**Received polarization:** left circular.  
**Receiver noise bandwidth:** 2.3, 11.0, 38.4 kc/s.  
**System losses:** 2 db.

Continuous calibration of the receiver was accomplished by injecting a pulse of wide-band noise of known intensity through a directional coupler at the receiver input. The calibrating pulse was injected near the end of the interpulse period to avoid interference with the backscatter returns.

Square-law detection was used for receiving. A digital encoder sampled the receiver output level at intervals of 250 microseconds. The digitalized output from the receiver was processed by a high-speed real-time computer which is part of the Millstone radar system. Computer-integration techniques were used to improve the apparent signal-to-noise ratio.

**Results.** The backscatter returns resemble noise when viewed on an A-scope display of the receiver output. Figure 1 is a photograph of such a display showing ground clutter followed by a rather extended-in-range noise signal.

Lowering the antenna elevation angle from the vertical causes the radar range to the backscatter



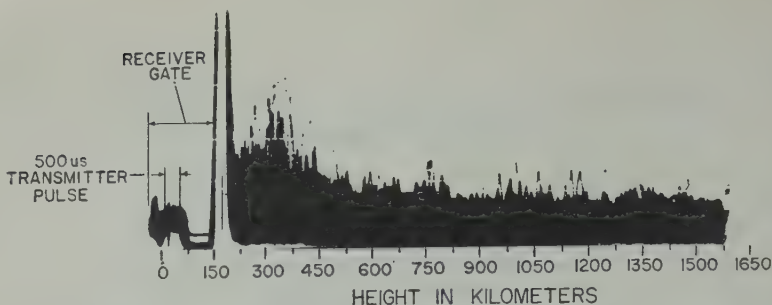


Fig. 1. Vertical-incidence backscatter at 440 Mc/s observed at 1434 hr EST, February 8, 1960.

ter to increase and the signal intensity to decrease. The general characteristics of the signal do not change with azimuth. There can be little doubt, therefore, that the observed returns are from the ionosphere.

*F*-region backscatter observations can be made with the antenna beam vertical, but weak returns from ground clutter and aircraft out to ranges of about 150 km usually obscure returns from the *E*-region. Consequently, it is necessary to operate with the antenna beam at a fairly low elevation angle in order to examine *E*-region backscatter.

The diurnal characteristics of ionospherically backscattered returns were observed during one 24-hour period in March 1960. Figure 2 shows the variation in height of maximum *F*-region signal intensities during this period.

Frequency spectrum measurements of ionospheric backscatter were made by sampling the received signal intensity at a number of frequencies in a band about 40 kc/s wide.

*F*-region frequency spectrums observed at a vertical-incidence height of about 300 km during February, March, and April tended to be flat-topped and were about 11 kc/s wide at the half-power points. The results of daytime and night-

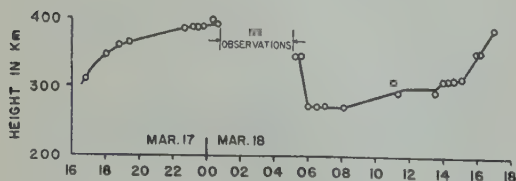


Fig. 2. Diurnal variation in height of maximum *F*-region vertical-incidence backscatter at 440 Mc/s. (The abscissa represents Eastern Standard Time).

time measurements of the *F*-region frequency spectrum made in March 1960 are shown in Figures 3 and 4. A similar spectrum measurement made in April 1960 with more attention to the shape in the region near the center frequency is shown in Figure 5.

The spectrum shown in Figure 5 is about 100 kc/s wide at the half-power points and shows a slight dip at the center frequency. Because of interference from ground clutter, it was necessary to lower the antenna beam to an elevation angle of 45° while taking the data for Figure 5. Despite the difference in antenna attitude, however, the difference between the spectral shape shown in Figure 5 and that shown in Figure 3 is believed to be the result of the more detailed observations in the central frequency region of Figure 5 rather than of any peculiar aspect phenomenon.

As was mentioned before, backscatter from the *E*-region of the ionosphere can be observed with the antenna beam at a low elevation angle. With this technique it has been possible to make simultaneous observations of the frequency spectrums of backscattered returns from both the *E* and the *F* region of the ionosphere.

The results of these simultaneous measurements, shown in Figure 6, indicate half-power band widths of 5 and 11 kc/s, respectively, for the *E*-region and the *F*-region frequency spectrums.

If, as is thought to be the case, the Doppler broadened frequency spectrums of both the *E* and the *F* region result from the thermal velocities of scatterers of the same kind, a direct comparison of the temperatures in the two regions is possible. The Doppler broadening is proportional to the square root of the kinetic temperature.

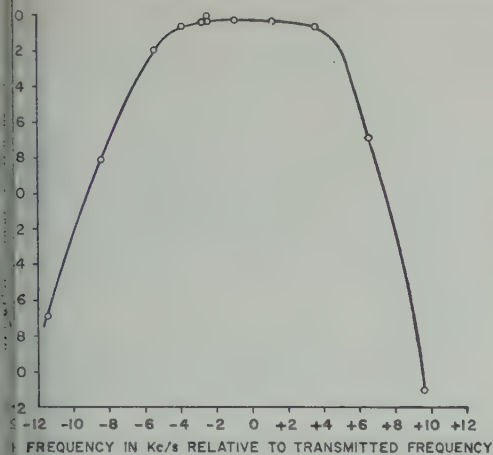


Fig. 3. Frequency spectrum of vertical-incidence backscatter from a height of about 300 km observed at 440 Mc/s between 1400 and 1500 hr EST, March 18, 1960.

Consequently, the ratio of the simultaneously observed frequency spectrums shown in Figure 5 indicates an *F*-region temperature only 5 times that of the *E*-region.

It should be noted that all the frequency spectrums shown in this paper are apparent spectrums. They do not take into account the effects of the transmitting and receiving bandwidths. These effects tend to cause the observed spectrums to be wider than the actual spectrums. The effect of this distortion would cause the ob-

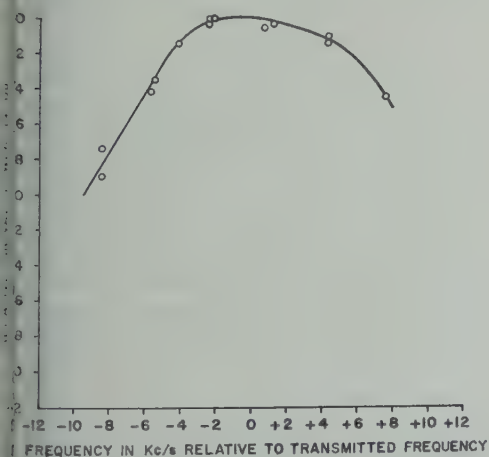


Fig. 4. Frequency spectrum of vertical-incidence backscatter from a height of about 380 km observed at 440 Mc/s between 2300 and 2400 hr EST, March 17, 1960.

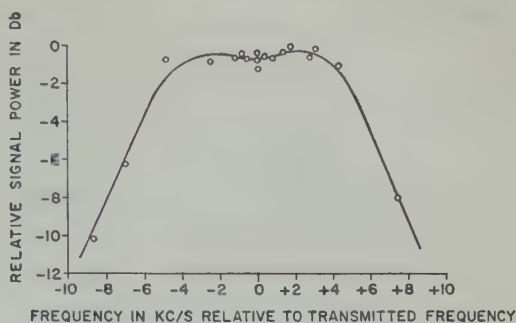


Fig. 5. Frequency spectrum of ionospheric backscatter observed at 440 Mc/s from a height of about 315 km between 1423 and 1515 hr EST, April 22, 1960. Antenna elevation = 45°; antenna azimuth = 281°.

served ratio of *F*-region to *E*-region spectrums to be less than the actual ratio.

A preliminary examination of the spectrum data taken at heights of 200, 300, and 600 km indicates a tendency for the half-power spectrum width to increase with height, although the change in width seems to be small in this height range.

No difference has been noticed between the day and the night *F*-region spectrum width.

The scattering coefficient for the ionosphere can be computed by means of the radar equation for volume-distributed scatterers [see Kerr, 1951].

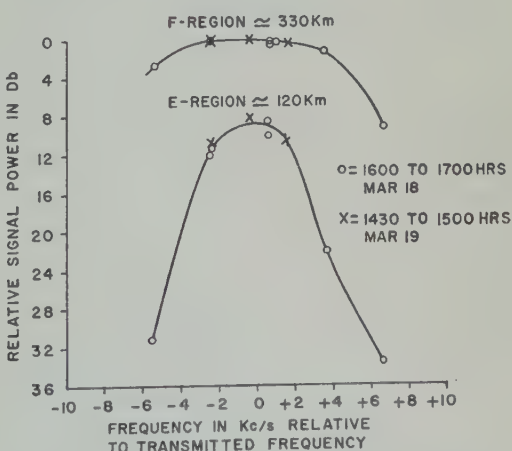


Fig. 6. Frequency spectrum of oblique-incidence backscatter received simultaneously from *E*-region and *F*-region heights at 440 Mc/s.

$$\overline{P_r} = (P_t A / 4\pi R^2)(c\tau/2)(1/L)^{N\bar{\sigma}} \quad (1)$$

where

- $\overline{P_r}$  = average receiver power.
- $P_t$  = peak transmitted power.
- $A$  = effective area of the antenna.
- $\tau$  = pulse length in seconds.
- $c$  = velocity of light.
- $N\bar{\sigma}$  = scattering coefficient.
- $N$  = number of scatterers per unit volume.
- $R$  = distance to the scattering volume.
- $\bar{\sigma}$  = average cross section of scatterers.
- $L$  = loss factor in transmission lines.

It is apparent from equation 1 that height profiles of electron density can be calculated from backscatter signal power measurement either if the value of  $\bar{\sigma}$  is known, or if the value for  $N$  at any particular height can be determined by some other method.

Electron densities at the height of  $F_2$  maximum were deduced from  $F_2$  critical frequencies measured at the AF Command and Control Development Division's ionosonde site in Bedford, Massachusetts, about 10 miles from Westford.

A typical height profile of electron density obtained by this method is shown in Figure 7.

Given the value of  $N$  corresponding to a particular backscatter power measurement, it would seem to be quite simple to determine the value of  $\bar{\sigma}$ . Here one must proceed with caution, however, for, although all the electrons in the scattering volume contribute to the power

available at the antenna terminals, only the electrons whose Doppler-shifted frequencies are within the receiver pass band can contribute to the receiver output power. Thus, in order to obtain a sensible value for  $\bar{\sigma}$ , it is necessary either that the receiver pass band be as wide as the bandwidth of the backscatter or that the ratio of the receiver bandwidth to that of the backscatter be known and taken into account to determine the power available at the antenna terminals.

Apparent values of the scattering coefficient  $N\bar{\sigma}$  of  $3.3 \times 10^{-18}$ ,  $10.4 \times 10^{-18}$ , and  $13.6 \times 10^{-18}$  were calculated, using equation 1, from backscatter signal intensity measurements in receiver noise bandwidths of 2.3, 11.0, and 38.4 kc/s, respectively. The corresponding values for  $\bar{\sigma}$  are  $2.1 \times 10^{-30}$ ,  $6.5 \times 10^{-30}$ , and  $8.5 \times 10^{-30}$  square meter, respectively. For reasons stated before, the value of  $\bar{\sigma}$  obtained with the largest receiver bandwidth is considered to be closest to the actual value.

The accuracy of the various measurements for obtaining the data used in calculating the values of  $\bar{\sigma}$  was about  $\pm 1.5$  decibels. This corresponds to an accuracy of about  $\pm 2 \times 10^{-30}$  square meter in the values for  $\bar{\sigma}$  obtained with receiver bandwidths of 11 and 38.4 kc/s.

**Conclusions.** The experiment described in this paper demonstrates the feasibility of using a high-powered UHF radar in conjunction with a high-speed digital computer to obtain information about the electron density and other characteristics of the ionosphere.

The backscattered signal is incoherent and has a Doppler-broadened frequency spectrum whose width increases with height. The observed half-power width of the  $F$ -region spectrum at the time and place of these observations was about 11 kc/s at a height of about 300 km. The observed half-power width of the  $E$ -region spectrum was 5 kc/s. Because of the effects of the transmitting and receiving bandwidths, the actual spectrum widths may be less than the observed ones.

The values for the scattering cross section obtained in this experiment are limited by the accuracies of the various measurements of the data from which they were derived, but a value between about  $5 \times 10^{-30}$  and  $8 \times 10^{-30}$  square meter is indicated. This is reasonably close

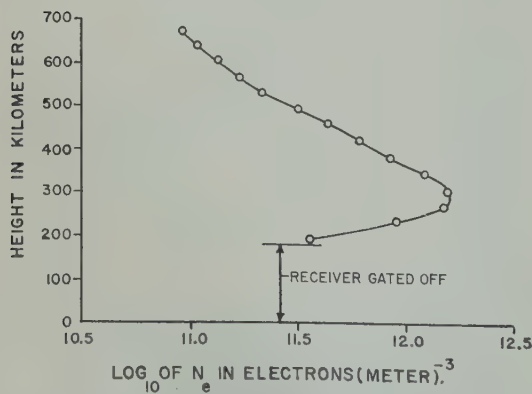


Fig. 7. Height profile of electron density obtained from vertical-incidence backscatter observations at 440 Mc/s at 1420 hr, EST March 18, 1960.



square of the classical electron radius,  
otherwise known as the Thomson cross section.

## REFERENCES

- s, K., Observation of vertical incidence  
scatter from the ionosphere at 41 Mc/s, *Phys.*  
*Letters*, *1*, 454-455, 1958.
- D. E., *Propagation of Short Radio Waves*,  
McGraw-Hill Book Co., New York, chapter 7,  
1951.
- Gordon, W. E., Incoherent scattering of radio  
waves by free electrons with applications to  
space explorations by radar, *Proc. IRE*, *46*, 1824-  
1829, 1958.
- Pineo, V. C., L. G. Kraft, and H. W. Briscoe,  
Ionospheric backscatter observation at 440 Mc/s,  
*J. Geophys. Research*, *65*, 1620-1621, 1960.



# Radio-Wave Scattering by an Ionized Gas in Thermal Equilibrium<sup>1</sup>

J. A. FEJER

*Theoretical Studies Group  
Defence Research Board, Ottawa, Canada*

**Abstract.** Expressions for the power density and the power spectrum of radio waves scattered by the density fluctuations of an ionized gas in thermal equilibrium are presented. These expressions are applicable to the terrestrial ionosphere and are in agreement with results obtained by an UHF radar.

Gordon [1958] has recently developed a theory of the scattering of radio waves by the density fluctuations existing in an ionized gas in thermal equilibrium and has suggested that this scattering may be detected by a powerful radar. Gordon [1958], by means of such a radar beamed vertically upward, has observed weak noise-like signals from ionospheric heights, of about the predicted power density but with a very much lower spectrum than had been anticipated by Gordon.

Gordon's predictions were made on the basis of scattering theory that neglects the Coulomb forces acting between the particles that constitute an ionized gas. If these forces are considered, it is found [Fejer, 1960] that the scattering coefficient  $\sigma$  per unit volume is given by

$$\sigma = \sigma_e N \left\{ \frac{[4\pi l_D \sin \theta/2]^2 + \lambda^2}{[4\pi l_D \sin \theta/2]^2 + 2\lambda^2} \right\} \quad (1)$$

where  $\sigma_e$  = scattering cross section of a single electron.

$N$  = electron number-density.

$l_D$  = Debye length.

$\theta$  = scattering angle.

$\lambda$  = wavelength of incident radio wave.

For small values of  $\lambda$ , equation 1 reduces to Gordon's result

$$\sigma = \sigma_e N \quad (2)$$

If  $\lambda$  is large compared with  $l_D$ , as it usually is for currently available radars and the terrestrial ionosphere, equation 1 becomes

$$\sigma = \frac{1}{2} \sigma_e N \quad (3)$$

A similar result for more general conditions has been derived by Kahn [1959].

The form of the power spectrum of the scattered waves must be known if the temperature of the ionized gas is to be deduced from such observations. For a random spatial distribution of electrons, the spectrum would be gaussian in shape with a half-width proportional to the electron temperature. For low collision frequencies and wavelengths greater than the Debye length, the Coulomb interactions narrow the spectrum considerably. In this case, which usually applies in the ionosphere, the spectrum function has the form

$$S(\omega) = \frac{\exp(-\Delta\omega^2/\Omega_1^2)}{[1 - (\Delta\omega/\Omega_1) \exp(-\Delta\omega^2/\Omega_1^2) \int_0^{\Delta\omega/\Omega_1} \exp(u^2) du]^2 + \pi[(\Delta\omega/\Omega_1) \exp(-\Delta\omega^2/\Omega_1^2)]^2} \quad (4)$$

This paper was presented at the Symposium on the Exosphere and Upper  $F$  Region sponsored by Commission III of the United States National Committee of the International Scientific Radio Union (ISRI), held on May 4, 1960, in Washington, D. C.



where

$$\Omega_1 = [8KT/m]^{1/2}(\omega_0/c) \sin \theta/2.$$

$T$  = temperature (assumed equal for electrons and ions).

$m$  = mass of positive ion.

$\omega_0$  = frequency of the incident wave.

$K$  = Boltzmann's constant.

This function is flat-topped, with two slightly raised shoulders situated symmetrically above and below the frequency of the incident wave. J. P. Dougherty and D. T. Farley (private communication, 1960) and E. Salpeter (private communication, 1960) by very different methods, have arrived at essentially the same expression. Experimentally this spectrum has been confirmed by results obtained at Lincoln Laboratory by Pineo, Kraft, and Briscoe [1960] for echoes returned from the  $E$  and  $F$  regions of the ionosphere. On the basis of the above expression these observations yield an  $F$ -region temperature of about  $1400^\circ$  and an  $E$ -region temperature of about  $400^\circ$ , on the assumption that atomic oxygen provides the dominant heavy ions.

Power spectra for the case of a collision-domi-

nated gas have also been derived and are presented elsewhere [Fejer, 1960].<sup>2</sup>

#### REFERENCES

- Bowles, K. L., Observations of vertical incidence scatter from the ionosphere at 41 Mc/sec, *Rev. Letters*, **1**, 454-455, 1958.
- Fejer, J. A., Scattering of radio waves by ionized gas in thermal equilibrium, *Can. J. Phys.*, **38**, 1114-1133, 1960.
- Gordon, W. E., Incoherent scattering of radio waves by free electrons with applications to space exploration by radar, *Proc. IRE*, **46**, 1829, 1958.
- Kahn, F. D., Long-range interactions in ionized gases in thermal equilibrium, *Astrophys. J.*, **129**, 205-216, 1959.
- Pineo, V. C., L. G. Kraft, and H. W. Briscoe, Some characteristics of ionospheric backscatter at 41 Mc/s, *J. Geophys. Research*, **65**, 2629-2633, 1960.

---

<sup>2</sup> Note added in proof. A more detailed account of this investigation is contained in a paper appearing in the August 1960 issue of *Canadian Journal of Physics* [Fejer, 1960]. The error mentioned in a note added in proof to that paper has been corrected in equation 1 above.

## Low-Energy Solar Cosmic Rays and the Geomagnetic Storm of May 12, 1959

J. R. WINCKLER AND P. D. BHAVSAR

*University of Minnesota,  
Minneapolis, Minnesota*

**Abstract.** During a strong geomagnetic storm on May 11–12, 1959, and following a class 3+ chromospheric solar flare on May 10, a large intensity of low-energy solar cosmic-ray protons was observed in a balloon flight at 30-km altitude at Minneapolis, Minnesota. This paper reports the results obtained with an ion chamber, Geiger counter, and scintillation counter. The proton flux of energy above 100 Mev varied from 52 to 1.7/cm<sup>2</sup>·sec·steradian during the time 0400 to 1800 UT on May 12. The spectrum obtained with the counting instruments is consistent with detailed emulsion studies and has the form

$$N(>E) = CE^{-\gamma}.$$

With  $E$  in Mev,  $\gamma$  varies from 5 to 2 during the above time period. The scintillation counter showed excess  $\gamma$  radiation from the nuclear effects of the cosmic-ray protons, and intense bursts of auroral X rays, particularly at 0835 UT on May 12. The solar optical and radio features of the flare are discussed, and the relationship of the polar-cap ionospheric blackout to the Minneapolis measurements is considered. The cosmic-ray protons at Minneapolis were measured at energies well below the normal Störmer cutoff for that latitude. Several theories on this subject are presented.

**Introduction.** A series of balloon flights was made at Minneapolis, Minnesota, during the geomagnetic storm of May 11–12 following the solar-flare event and cosmic-ray acceleration of May 10, 1959. In this paper we will discuss mainly a single long flight of this series using an integrating ion chamber, a Geiger counter, a scintillation counter, and nuclear emulsions, covering the period from 0400 UT on May 12 to 0100 UT on May 13. We shall be concerned in this paper with the time variation, photon fluxes, and the ionization and counting effects produced by the solar cosmic rays and the auroral X rays. The correlation of other cosmic-ray and geophysical data is of importance and will also be considered. The analysis of the data during this flight series will be presented in a separate communication by E. P. Winckler and others. A preliminary communication [Winckler, and Freier, 1959] has discussed the main features observed at Minne-

apolis and about 20°N latitude at 2055 UT on May 10. The flare was accompanied by extremely intense solar radio noise bursts classified as type IV or wide-band. This solar radio noise storm lasted from 2118 UT on May 10 to 0200 UT on May 11. The terrestrial effects of the flare included the usual prompt ionospheric electron-density increase which was observed, for example, as a sudden cosmic-noise absorption in the polar regions resulting from the ultraviolet or soft X-ray fluxes from the flare. The first detection of cosmic-ray particles arriving over the polar region was reported by Reid and Leinbach [1959], beginning about 0130 UT on May 11 but continuing to increase in intensity until 1600 UT on May 11. These low-energy solar cosmic rays were evidenced by the enormous increase in the  $D$ -layer absorption of the cosmic radio noise received at College, Alaska, and at many other stations in the polar region, and by a complete solar blackout of short-wave radio communication.

Direct evidence for the early arrival of the cosmic-ray particles at high latitude has also been obtained by Charakhchian, Tulinov, and Charakhchian, [1960] with a cosmic-ray balloon

flight made on May 11 at 1000 UT which showed a considerable increase in the counting rate of a single Geiger counter at high altitude above the normal cosmic-ray background. Subsequent flights made from the same station at Murmansk at a geomagnetic latitude of  $64^\circ$  confirmed the continued incidence of the cosmic-ray particles and gave information about the energy spectrum similar to what will be described in this paper. The next definite terrestrial consequence of the flare was the sudden commencement of the magnetic storm which occurred at about 2320 UT on May 11. This coincided with the beginning of a decrease of the galactic cosmic radiation (Forbush decrease) as shown by sea-level neutron monitors on a world-wide basis. The decrease in galactic cosmic rays continued until about 0400 UT on May 12, when the bottom of the Forbush decrease was reached. The magnetic storm was of very high intensity, with planetary  $K$  indices reaching 9 for approximately 12 hours. This storm subsided by 2000 UT on May 12. The balloon flights at Minneapolis all occurred during the intense parts of the geomagnetic disturbance, and it is clear that the low-energy cosmic rays appeared at Minneapolis only during the terrestrial magnetic storm. It is also certain that a continuous flux of these particles arrived at the polar cap, beginning a few hours after the flare, and continued in an uninterrupted but slowly decaying stream for more than 5 days. There appears to be no good evidence for any trapping of the cosmic rays in the low-energy plasma cloud from the flare. The geomagnetic storm was accompanied by an intense aurora which was observed at the zenith at Minneapolis, a great auroral maximum being reached at 0835 UT on May 12. It is noteworthy that, after the solar cosmic-ray increase from the big flare on May 10 had already begun, a second smaller flare occurred on May 11 at 2006 UT with coordinates east 39, north 08, classed as 2+. This flare also produced a short burst of continuum radio emission. There is no certain evidence that this smaller flare accelerated low-energy cosmic rays.

*Solar and ionospheric details of the May 10-12 event, 1959.* The chromospheric flare which is the almost certain source of the low-energy cosmic rays has been described as the largest of the present solar cycle. Table 1 gives relevant

TABLE 1. Summary, Flare of May 10, 1959

|                    |                          |
|--------------------|--------------------------|
| Start              | 2055 UT                  |
| Maximum phase      | 2115-2150 UT             |
| End                | 2600 UT (0200 UT 11 May) |
| Coordinates        | N19 E46                  |
| Plage region       | McMath no. 5148          |
| Ionospheric effect | Slow S-SWF               |
| Time: 2110         |                          |
| Duration: 500 min  |                          |
| Importance: 3+     |                          |

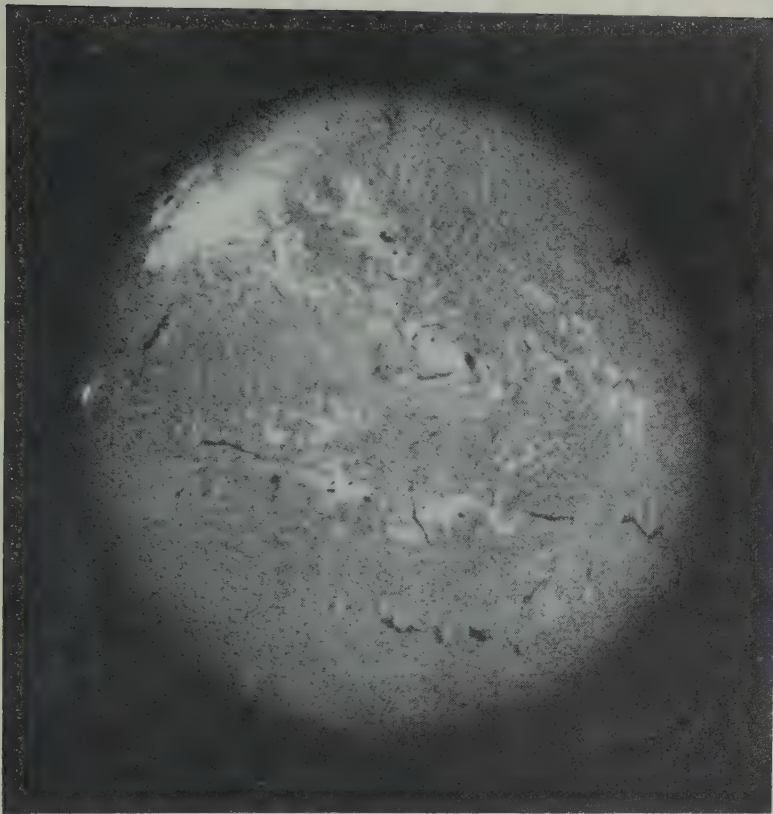
\* From Solar-Geophysical Data, CRPL-FF Part B, National Bureau of Standards, Boulder, Colorado.

details of this eruption. The area of the flare extremely large, as can be seen, amounting over 100 square degrees on the sun. In Figure 2 we show a photograph of the flare in the  $H_\alpha$  through the courtesy of the Sacramento Peak Observatory, Sunspot, New Mexico. Actually this main, class 3+, flare was followed almost immediately by a second, class 3, flare seen at Sydney and Mitaka from 2315 to 2610 UT at the same location. It is evidently part of the same flare complex and not necessarily a separate phenomenon.

The flare emitted bursts of radio-frequency continuum radiation of very high intensity. Figure 2 shows the radio output of the flare at a number of frequencies (A. D. Fokker and L. de Feiter, Netherlands PTT, The Hague, private communication; R. T. Hansen, High Altitude Observatory, U. S. Bureau of Standards, Boulder Colorado, private communication). These radio records display the initial bursts between 2120 and 2150 UT, many features which occur on all frequencies from 18 to 100 Mc/s. The noise continuum beginning at about 2200 UT on May 10 and persisting during a part of May 11 is also of wide-band character. There is a possible close correlation between these long-lasting type IV increases and the cosmic-ray acceleration process. Both necessitate solar magnetic fields, and the type IV emission is due to the synchrotron process, necessitates energetic electrons that can be accelerated along with the protons.

The region containing the flare was active in the emission of radio waves all through the traversal of the sun. In Figure 3 we reproduce a measurement due to Erickson [1959] v





The great flare of May 10, 1959, taken in the light of  $H_{\alpha}$ . (Photograph courtesy of Sacramento Peak Observatory, Air Force Cambridge Research Center, Sunspot, New Mexico.)

of the apparent position of a source of radio at 26.3 Mc/s which is associated with the McMath no. 5148 containing the flare on May 10. Erickson plots the apparent position from radio interferometer observations and compares the expected position, considering refraction in the solar corona, for a region 4, 4.5, 5 solar radii from center. The observations are compared with the computed transit times and show that the region may be as much as 4, 4.5, 5 solar radii from center. The plasma frequency is about 26 ms/s at about 1.6 solar radii, so the observed emission at 4 or 5 solar radii must be from plasma oscillations in the corona. The region apparently rotated with the sun to a position in the solar corona approximately 5 solar radii from the center.

These regions have been described also by [1959] and co-workers, and the radiation has been attributed to synchrotron radiation of electrons in the solar magnetic field. More

recently, *Denisse* [1960] has proposed that the radiation may be Cerenkov radiation from electrons falling back into a magnetic plasma region in the upper corona. The Cerenkov radiation would be caused by electrons, and also possibly by the solar protons, if the velocity is greater than the velocity of propagation of plasma waves in the solar medium. The interest here is that such magnetic 'bottles' may be closely connected with the production or storage of the solar cosmic radiation which can be injected by the chromospheric flare below. The electrons responsible for the radio emission would lose energy rapidly by radiation, but the protons would remain for much longer periods. It is of great interest that this 'hot' region at 5 solar radii increased its radio output after the great flare at 2100 UT on May 10, shown in Figure 1. A coupling between the flare region and the radio-emitting region by either magnetic fields or particles or both, must be envisioned.

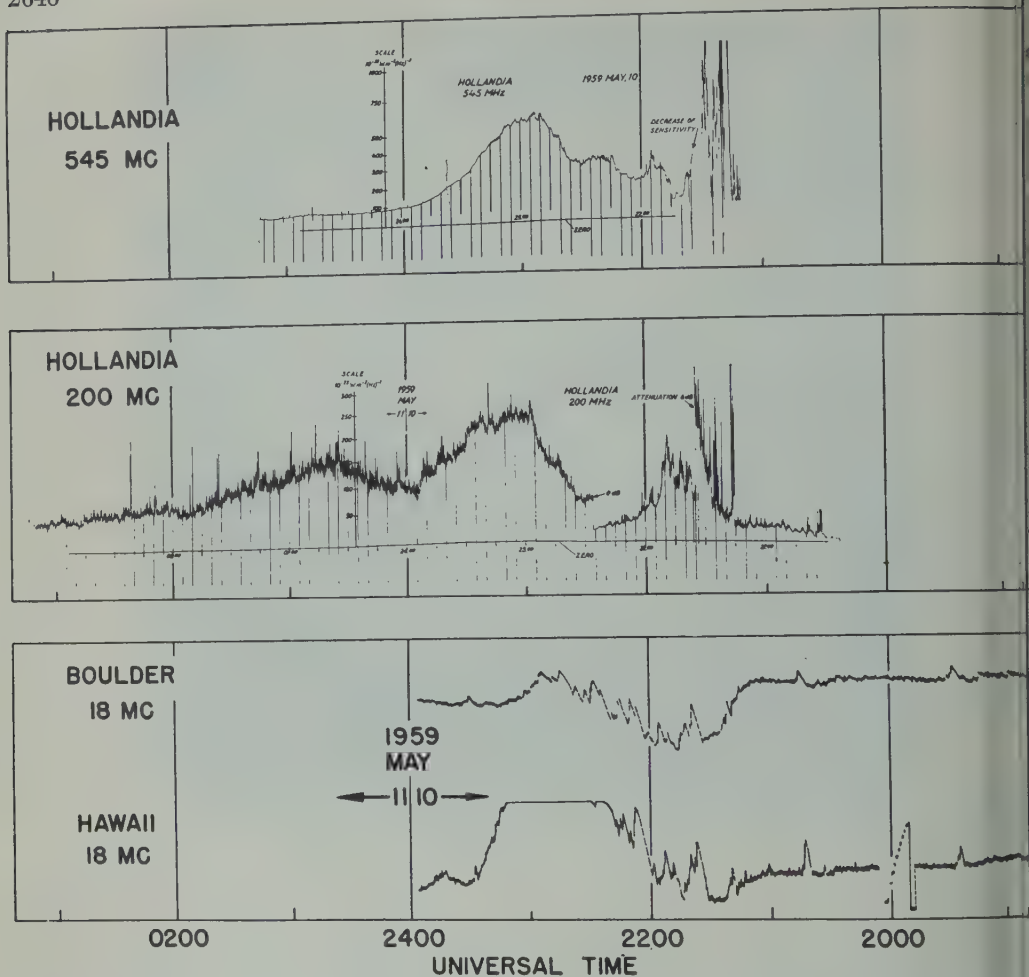


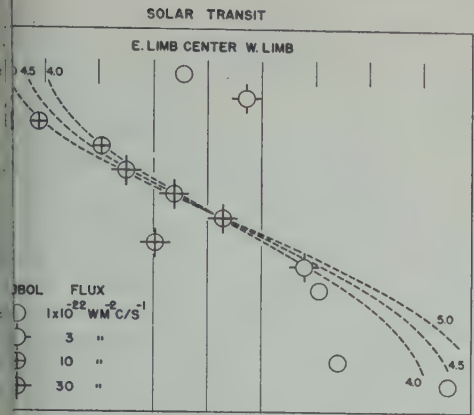
Fig. 2. Radio bursts from the flare of May 10, 1959, on a number of frequencies and locations. Upper observations at Hollandia, courtesy of A. D. Fokker; lower observations at Boulder, Colorado, and Hawaii, courtesy of Boulder Laboratories, U. S. Bureau of Standards, and High Altitude Observatory. (R. T. Hansen).

*Detailed experimental results of flight IGC-8.* The scintillation counter, ionization chamber, and Geiger counter can be interpreted to give information about the photon and particle fluxes and particle ionization and energies during this event. The total ionization in excess of normal cosmic rays as a function of atmospheric depth as the balloon rises into the flux of solar protons may be unfolded in the manner described by Anderson, Arnoldy, Hoffman, Peterson, and Winckler, [1959] to give the energy spectrum of the particles if their nature is known. Nuclear emulsions flown on the same flight may be used

to check the validity of this method, which then be extended to cases where emulsion data are not available. For this reason we treat matter in some detail.

The Geiger counter gives an independent measure of the flux of particles from which energy spectrum may also be unfolded. The ionization chamber and counter should be individually self-consistent, and in addition the ionization per-count ratio should be proper at each depth for the spectrum derived from each device.

The scintillation counter, when corrected for the particle background including edge effects



3. Radio emission from the flare region during its transit of the solar disk, courtesy of J. Erickson. The apparent positions are shown by the observational points. The dotted curves show the expected positions as a function of  $r$ , assuming the emitting regions to be at 4, 5 and 5 solar radii from the center of the sun.

the photon intensity. Its efficiency for photons is much higher than that of either of the other instruments, approaching 100 per cent at energies of interest to this discussion. These photons arising from auroral X rays and from cosmic-ray nuclear collisions in the atmosphere and other causes can be studied as a function of depth. When the balloon reaches its floating altitude, the time variations of the photons and particles can be studied and the mean ionization measured by comparing the ionization chamber and the counter. In balloon flights dur-

ing very strong storms, bursts of auroral X rays are frequently seen; they occurred on flight IGC-8.

To evaluate the data, the instrumental characteristics must be known in detail. Such a summary is given in Tables 2a, b, and c. Descriptions of these instruments and the details of the electronics, etc., have been published [Winckler, Peterson, Arnoldy, and Hoffman, 1958; Winckler, Peterson, and Howard, 1958; Winckler, 1960]. Besides the absolute size and sensitivity, an important feature is the ratio of the normalized ion chamber rate to the normalized single counter rate. This ratio is a measure of the mean ionization of the particles present on an omnidirectional basis and can be compared with normal cosmic rays at various depths from the table. The value for pure minimum particles must be computed. This can be done quite well; the normalized ratio is  $3.19 \times 10^{-4}$ . A frequently used source of laboratory calibration is  $\text{Co}^{60}$   $\gamma$  rays, and if the ion chamber and counter are exposed to a uniform flux of these  $\gamma$  rays the ratio  $4.03 \times 10^{-4}$  is obtained. This ratio, of course, applies only to the specific wall thicknesses, etc., used in the particular instruments flown. The scintillation counter has been independently calibrated using fluorescent X rays of copper, molybdenum and silver,  $\text{Cs}^{137}$ , and  $\text{Co}^{57}$ , and has a good energy scale. The constant for obtaining the equivalent number of ion pairs produced per cubic centimeter per second of standard air from the ion chamber rate can

TABLE 2a. Characteristics of Counting Instruments on Flight IGC-8

| Instrument                | Wall Material and Thickness, mm           | Geometry and Dimensions, mm                                   | Omnidirectional Projected Area, cm <sup>2</sup> | Other Details  |
|---------------------------|---|---|---|--|
| Ion chamber               | Steel<br>0.46                             | Sphere<br>Diameter = 254                                      | 506   | Neher-type pulsing chamber<br>$1.90 \times 10^{-10}$ coulomb per pulse<br>Filling gas: argon at 7.83 atm |
| Counter                   | Brass<br>End 1.32<br>Side 0.79<br>End 4.2 | Cylindrical<br>Active length = 38.0<br>Active diameter = 49.2 | 24.17   | Supplied by Wood Labs, self-quenching<br>Dead Time $\approx 200 \mu\text{sec}$                           |
| Scintillator<br>NaI (Tl°) | Aluminum<br>1.47                          | Cylindrical<br>Height = 12.7<br>Diameter = 45.45              | 12.64   | Harshaw scintillation crystal  |



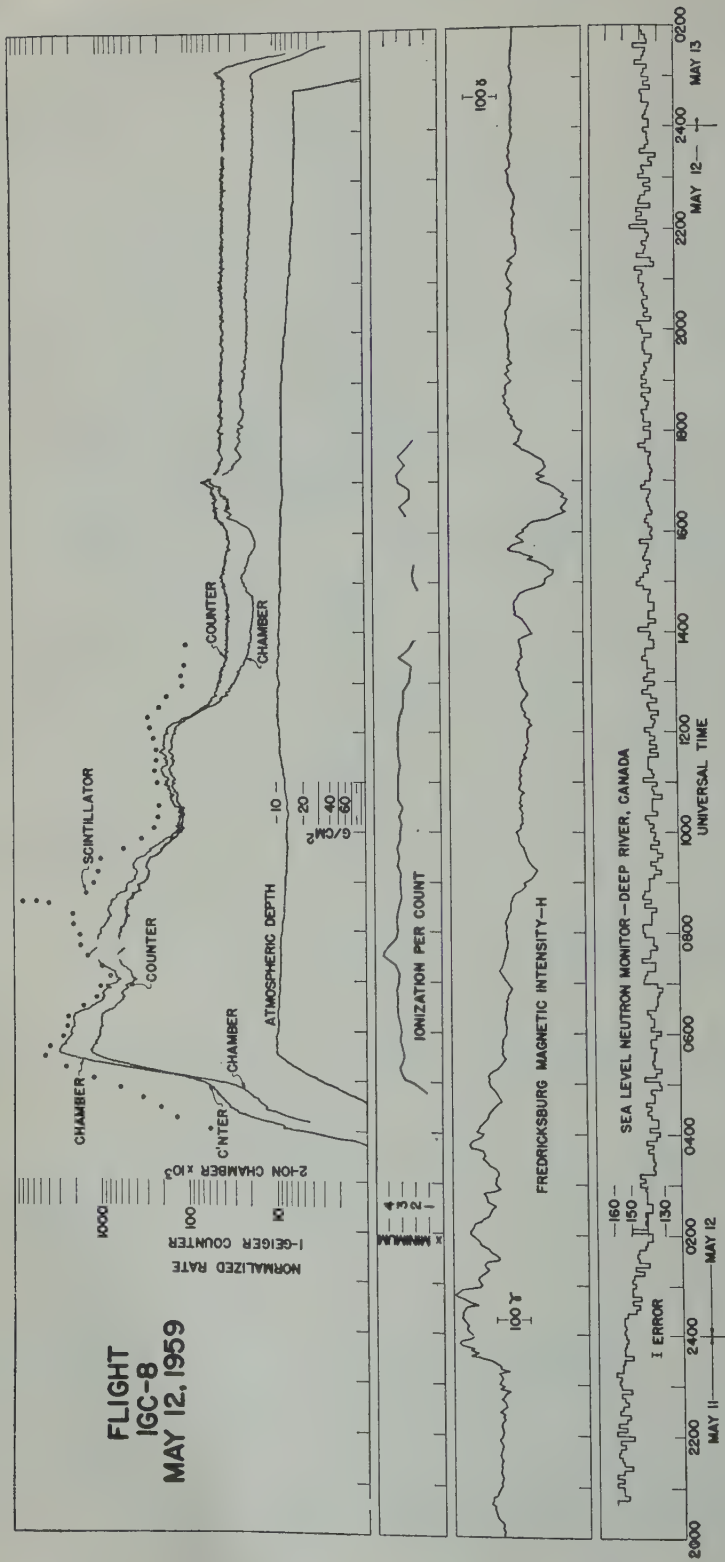


Fig. 4. Summary of information concerning the solar cosmic-ray increase observed at Minneapolis. The upper curve, flight IGC-8, shows the large increase of the Geiger counter, ion chamber, and scintillator after 0500 UT as the balloon rose into the low-energy proton radiation. Note the logarithmic scale. The atmospheric depth of the balloon expressed in  $\text{g}/\text{cm}^2$  is shown on the lower curve in the upper box. Below this is shown the ionization/count ratio, which is a measure of the mean energy of the particle spectrum; next below, the Fredricksburg magnetic intensity, showing the beginning of the storm at about 2320 UT on May 11 and the recovery about 2400 UT on May 12. The fluctuation at 1500 UT on May 12 may be the sudden commencement of another storm and is accompanied by changes in the ionization at high altitude. The Forbush decrease in sea-level neutrons is shown in the bottom scale; it began coincidentally with the beginning of the magnetic disturbance.

ated as follows:

pairs/cm<sup>2</sup>·sec·atmosphere air

24°C, 760 mm Hg)

= 12.3 × (normalized pulsing rate × 10<sup>3</sup>)

Figure 4 summarizes the readings of the instruments during the flight IGC-8. In the upper part of the figure is plotted the normalized rate of the ion chamber, the Geiger counter uncorrected for dead time, and the scintillation counter which has no dead-time effects. The next line down is the atmospheric depth of the balloon, which can be seen to remain between 10 and 15 g/cm<sup>2</sup> during the entire 21-hour flight. The next line below shows the ionization-per-count ratio of the ion chamber and counter during the flight. This value remains between 3 and 4 most of the time (a more detailed graph of this quantity is given in Fig. 8). The Fredericksburg magnetic intensity is shown on the next line to display general features of the storm, although Fredericksburg is a mid-latitude location and is not necessarily a very good example of the world-wide storm. The sudden commencement was at 01 UT on May 11; the positive phase lasted about 6 hours, and then the storm went into a negative phase with several bay disturbances. The average negative phase of this storm was considered considering the high *K* indices. In the bottom line, we plot the Deep River, Canada, neutron monitor (courtesy of Hugh Carmichael, High Data Center A for Cosmic Rays), which shows the considerable Forbush decrease beginning at the time of the sudden commencement and reaching its minimum intensity at about 0500 UT on May 12.

Flight IGC-8 left the ground at 0340 UT, and the first evidence for a large additional flux was when the balloon reached about 80 g/cm<sup>2</sup>. As can be seen, the rate curves of the ion chamber and counter began to bend over as the balloon passed through the usual Pfotzer maximum, and then with further decreasing atmospheric depth the rates increased to a very high value, more than two orders of magnitude above normal until the balloon reached ceiling altitude at 05 UT. The balloon then floated level, but the rates in general decreased with time, though with major fluctuations associated in some way with the features of the magnetic storm.

We turn our attention first to the ascending part of the flight. A detail of this is shown in

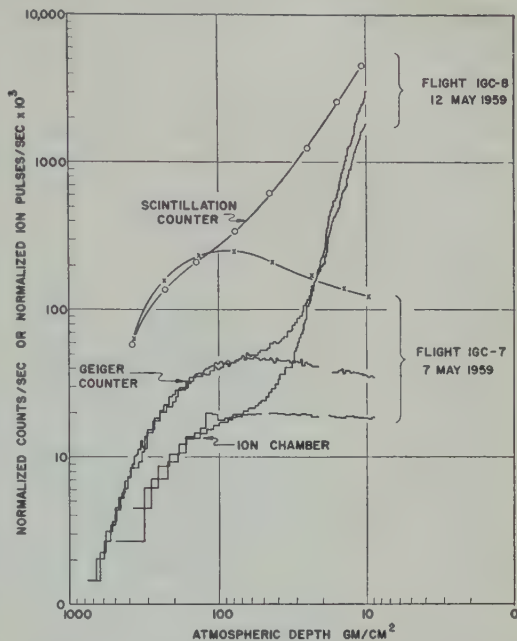


Fig. 5. Ion chamber, Geiger counter, and scintillation counter rates during ascent of the balloon on May 12, 1959, showing the comparison with a 'normal day' flight, IGC-7, about 1 week before the storm. Note that the first detectable effect of the solar protons occurs at about 80 g/cm<sup>2</sup> at atmospheric depth on the ion chamber and counter. The scintillator detects an effect at lower altitude, perhaps to 150 g/cm<sup>2</sup>, probably due to  $\gamma$  rays. The instrument rates were still rising steeply when the balloon reached its ceiling altitude of about 10 g/cm<sup>2</sup>.

Figure 5, in which the normalized rates of the scintillator, counter, and ion chamber are plotted. This is compared with flight IGC-7, made about a week earlier under approximately normal cosmic-ray conditions. It can be seen that the rate increases continually to the highest altitude reached, 10 g/cm<sup>2</sup>. This means that the energy spectrum of the incident particles, assuming protons, goes up to at least 100 Mev. The immediate conclusion is that particles are incident at this time well below any previously assumed geomagnetic cutoff energies for the latitude of Minneapolis. Recent estimates place this value between 300 and 400 Mev for protons [McDonald, 1957; McDonald and Webber, 1959; Fowler, Freier, and Ney, 1958]. For extracting the energy spectrum from the observed altitude dependence, we plot in Figure 6 the excess particles above normal cosmic-ray intensity, as ob-

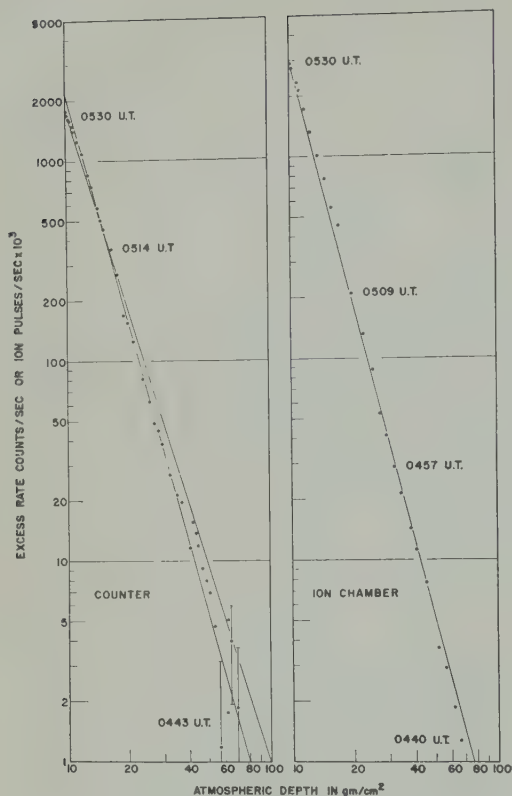


Fig. 6. Excess rates observed in the Geiger counter and ion chamber plotted as a function of atmospheric depth. The straight lines represent possible fits to the data (at two different ranges of atmospheric depth, in the case of the counter). The counter rate has been corrected for dead-time effects.

served in the counter and as also observed in the ion chamber, on log-log coordinates.

We consider first the Geiger counter curve. A simple interpretation of this curve may be made if we assume that the solar cosmic rays are all protons, arriving isotropically, and that their spectrum obeys a power law in which  $\gamma$  is the exponent of the differential energy spectrum having the form

$$N(E) = CE^{-\gamma} \quad (1)$$

The energy is related to the range of the particles according to a second-power-law relation,

$$R = K_1 E^\alpha$$

having the exponent  $\alpha$ , which, for protons in this

energy range, has a value of approximately 1. Then the number of protons crossing the counter at depth  $h$  from the vertical direction consists of those having range greater than  $h$ , which is given by the integral range spectrum of the form

$$N_1(>R) = KR^\beta$$

where

$$R = h,$$

$$\beta = (1 - \gamma)/\alpha,$$

$$K = C/(\gamma - 1)K_1^{-\beta}.$$

And therefore, by using the Gross transformation the omnidirectional intensity  $N_h$  at a depth  $h$  can be given by

$$N_h = 2\pi K/(1 - \beta) h^\beta$$

or

$$\log N_h = \beta \log h + K_2$$

where  $K_2 = \log 2\pi K/(1 - \beta)$ . The slope  $\beta$  of the logarithmic equation is thus a function of  $\gamma$  and  $\alpha$ , and from the measured slope of the counter curve, which is  $-3.5$ , we get the value  $\gamma = 7$  for the differential energy spectrum of the protons.

In considering the ionization chamber curve shown in Figure 6, unfortunately no such simple analysis is possible. However, the ion chamber results can be treated correctly by the method shown by Anderson, Arnoldy, Hoffman, Peterson, and Winckler [1959]. In Figure 7 we show curves computed using the complete analysis for a number of assumed exponents of the differential energy spectrum. The experimental points for the counter and ion chamber, when placed on these curves, show that a value of  $\gamma$  between 6 and 7 is in reasonable agreement with both sets of curves. In this analysis, the dead-time correction of the Geiger counter is important, and we have used a value of 100 microseconds, which was determined by measurements on a number of counters of this type. Nuclear interactions are not important if the atmospheric depth is not too great. There is no reason to believe that the measured value of  $\gamma$  is higher than the actual value, since presenting the spectrum of the particles as a function of that this is the result of a time variation of the primary beam. This will be discussed below in more detail.



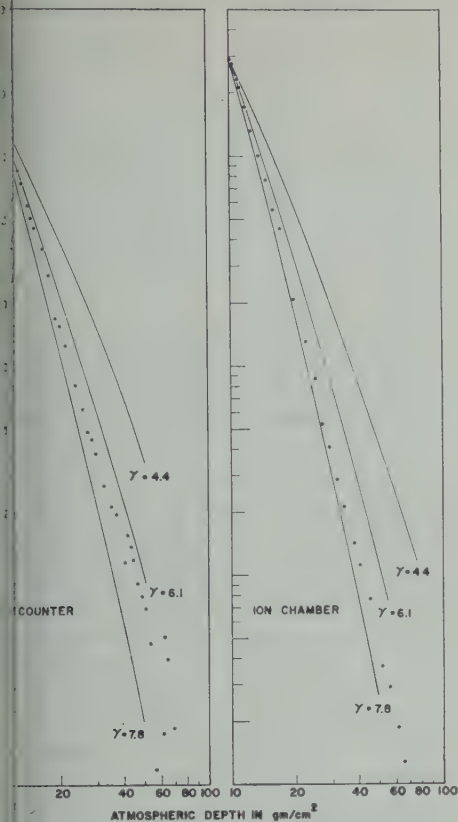


Fig. 7. Computed depth variation assuming values of the exponent  $\gamma$  of the power-law spectrum assumed for the incident solar protons. The value of  $\gamma$  somewhat in excess of 6 is in agreement with both curves. This value is thought to be somewhat higher than the correct one, however, due to a time variation that occurred during the flight of the balloon.

Now consider the constant-level part of the balloon flight. From the excess rates of the instruments at ceiling altitude, it is possible to determine the exponent of the energy spectrum of the protons from the observed mean ionization at the known depth of the balloon. In Figure 8 are shown the detailed relative ionization curve for the period of the flight in which detectable numbers of particles above background were encountered. The  $I/I_{min}$  values in the lower curve are those obtained after a careful dead-time correction for the Geiger counter, using the ratio for minimum ionizing particles computed for these instruments as described in Figure 2b. It is seen that the mean ionization

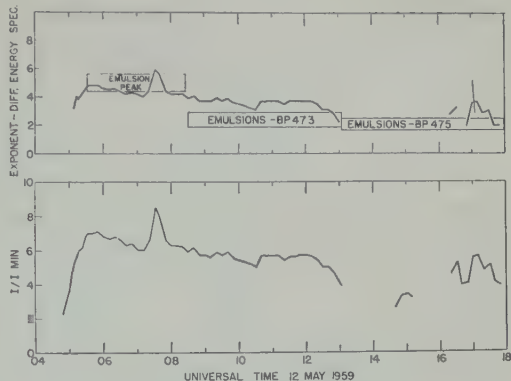


Fig. 8. Relative mean ionization of the excess radiation and the resultant exponent of the differential energy spectrum plotted as a function of time during the ceiling altitude of the balloon. Note that the radiation hardens as time proceeds, with one large increase at about 0800 UT. Good agreement is obtained with spectra measured in detail in emulsion flights flown during the period shown by the rectangles in the upper curves. The details of this separate measurement by Freier and Ney will be discussed in a separate communication.

climbs rapidly as the balloon descends into the proton flux and then, in general, descends during the rest of the geomagnetic storm with the exception of a considerable spike at about 0730 UT. In the upper broken curve we show the computed exponent of the differential kinetic energy spectrum, taking into account the observed depth of the balloon during its 'constant level' flight and assuming isotropic incidence at the top of the atmosphere and a power-law spectrum for the protons. In the same figure we show the results obtained with emulsion flights analyzed by Ney and Freier covering the same period. These emulsion results, which will be described in detail in a further paper of the Minnesota cosmic-ray group, are averages over the time intervals indicated. They are seen to be in reasonably good agreement with instrumental results. This good comparison with the unambiguous emulsion measurements increases our confidence that the instrumental analysis is essentially correct when all factors are properly taken into account. During the period that the spectrum flattens, of course, there is a large decrease in intensity, which is shown in Figure 4. The transitory increase of intensity at 0730

TABLE 2b. Values of Ion Pulses per Count (normalized ion chamber rate/normalized single counter rate) for Quiet-Day Cosmic Rays and  $\gamma$  Rays

| Source of Data   | Atmospheric Depth, g/cm <sup>2</sup> | Ratio $\times 10^4$ |
|--|--------------------------------------|---------------------|
| Observed   | 300                                  | 3.36                |
| Observed   | 200                                  | 3.84                |
| Observed   | 100                                  | 4.03                |
| Observed   | 50                                   | 4.21                |
| Observed   | 10                                   | 5.02                |
| Calculated for isotropic minimum ionizing particles    |                                      | 2.35                |
| Observed Co <sup>60</sup> $\gamma$ rays isotropic flux |                                      | 4.03                |

UT shown in Figure 4 is reflected in the exponent of the spectrum, as at this time the exponent rapidly climbs to the very high value of 6, then decreases, giving the more or less smooth trend shown in Figure 8, upper curve.

From the single counter rate  $N_h$  at a particular instant, and atmospheric depth  $h$ , and the knowledge of the exponent of the energy spectrum from ion chamber/single counter ratio, can also be calculated the integral intensity of the excess protons at the top of the atmosphere with energies greater than that required to penetrate down to  $h$ . These values are given in Table 3.

It seems clear that these changes in the shape of the spectrum are a property of the geomagnetic field of the earth acting on the solar proton beam as it exists in free space. As far as we can tell, the polar-cap effects continue without change, other than the steady decay during the period in which the geomagnetic storm produces the enormous increases observed at Minneapolis. From the Geiger counter and ionization chamber results, only mean values over the spectrum can be obtained and the presence or absence of sharp cutoffs cannot be determined,

TABLE 2c. Scintillation Counter Calculated Efficiencies for Omnidirectional X-Ray Flux

| Energy, kev | Efficiency, % |
|-------------|---------------|
| 50          | 100.0         |
| 100         | 99.99         |
| 250         | 69.90         |
| 500         | 41.70         |

TABLE 3. Proton Flux Values\*

|         | Time, UT  | $E$ , Mev | $\gamma$ | $C$                | $N(>E)$ sec $\cdot$ cm <sup>2</sup> |
|---------|-----------|-----------|----------|--------------------|-------------------------------------|
| Peak    | 0539      | 105       | 4.8      | $2.47 \times 10^9$ | 51.5                                |
| Average | 0539-0800 | 105       | 4.8      | $1.45 \times 10^9$ | 30.2                                |
| Average | 0800-0940 | 105       | 3.5      | $1.23 \times 10^9$ | 10.9                                |
| Average | 0940-1300 | 105       | 2.8      | $7.46 \times 10^8$ | 1.7                                 |

\* See equation 1 for definition of  $E$ ,  $\gamma$ , and  $C$

at least from the data obtained at constant  $E$  at high altitude. The existence, however, of considerable changes in the shape of the energy spectrum during the storm, when such changes are presumed not to exist in the solar beam outside the earth, presents an interesting problem in the action of the magnetic field of the earth during the perturbed condition of the storm on this solar beam. We hope to get some approximate idea about the cutoff energies during the storm from the intensity of the secondary  $\gamma$  rays detected by the scintillation counter. These  $\gamma$  rays are produced by the low-energy particles which themselves do not reach the balloon because of air cutoff.

A very important consideration for understanding the propagation of the solar cosmic rays through space is the observation that the particles are continuous at high latitudes, and the highly variable nature of the radiation seen at lower latitudes, for example at Minneapolis, is evidently controlled by local effects on the earth's magnetic field. The high-latitude radiation is of several kinds. On previous events *Anderson, Arnoldy, Hoffman, Peterson, and Winckler* [1959] have observed the radiation directly with balloons at Churchill and Fairbanks. On the present event, balloon observations at high latitude were made by *Charakhchian, Tubert, and Charakhchian* [1960] with cosmic-ray sounding balloons flown at geomagnetic latitude  $64^\circ$ . A number of solar cosmic-ray events have been observed with these sounding flights, including the great series in July 1959, on July 8, 1959, and also the event under present discussion covering the period May 11-15, 1959.

From Charakhchian's paper we reproduce in Figure 9 the sounding obtained on May 11, 1959, at 1000 UT with a single Geiger counter at

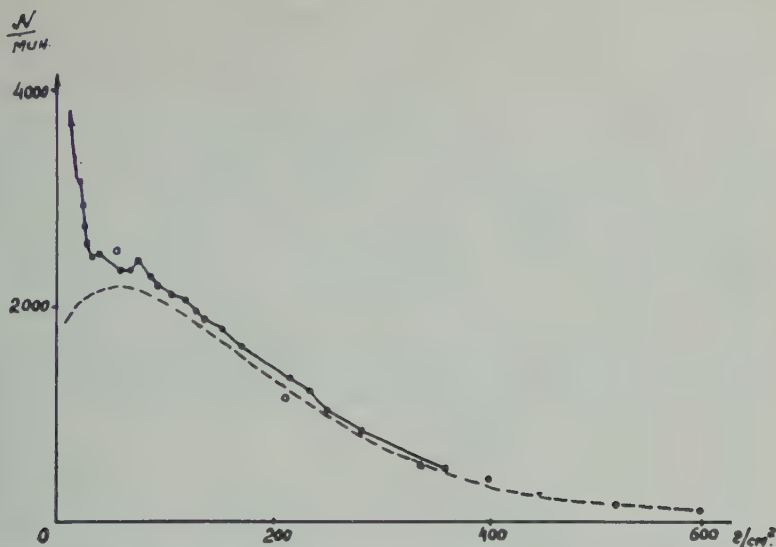


FIG. 10. Single-counter sounding made on May 11, 1959, at 1000 UT by Charakhchian and co-workers at magnetic latitude  $64^\circ$  in Murmansk, USSR. This shows the influx of solar cosmic rays at high altitude before the beginning of the geomagnetic disturbance which provided the entry of the particles to the latitude of Minneapolis. The dotted curve is for a sounding under normal conditions.

magnetic latitude. The curve shows the normalized cosmic-ray intensity by the dotted line and the high-altitude increase by the solid points. The latter shows a rapid rise below an atmospheric depth of about  $50 \text{ g/cm}^2$ . Besides the Geiger counter, vertical coincidence telescopes were also flown; these permitted the evaluation of the energy spectrum of the particles in a manner similar to that described in this paper. The energy spectrum showed a steeply rising with decreasing energy. Assuming a power-law spectrum, an exponent of 6.5 fits the Charakhchian measurements in May 12, 1959, on the same date as the Minneapolis measurements reported here. This value is somewhat higher than the values obtained with the constant-level part of the flight IGC-8 at Minneapolis but somewhat lower than the result obtained on ascent of flight IGC-8. We have assumed that the value of  $\gamma$  obtained on ascent of IGC-8, is high because of a time rate of increase of the primary rays at this time. However, the comparison with the value  $\gamma = 7$  obtained at Minneapolis and the value  $\gamma = 6$  obtained by Charakhchian may well be within the experimental errors of the methods. The solar cosmic

rays during this event were first observed by Charakhchian and co-workers on May 11 at 1000 UT; the event was observed until May 15. The event was not observed at Moscow, latitude  $51^\circ$  geomagnetic, or latitude  $41^\circ$  geomagnetic, although at these latter stations the Forbush decrease was noted on the sounding flights. No large effects are reported coincident with the geomagnetic storm beginning on May 12.

Besides these direct high-latitude measurements of the cosmic rays with a radiation detector in the high atmosphere, the ionospheric effects were very strong on this event over the polar regions. For example, the event was detected by the 'riometer' (cosmic noise absorption) network operated by Reid and Leinbach [1959] at the Geophysical Institute, University of Alaska. The ionospheric effects were also observed in an analysis of the ionospheric scatter network circuits over the arctic by Dana Bailey (private communication). In Figure 10 we show Bailey's results for the May event. In the scatter link, a slight ionization increases the signal strength and an additional ionization in the D layer of the ionosphere attenuates the signal. In Figure 10 the signal intensity over five different propagation circuits is shown. The position of these propagation links is outlined in



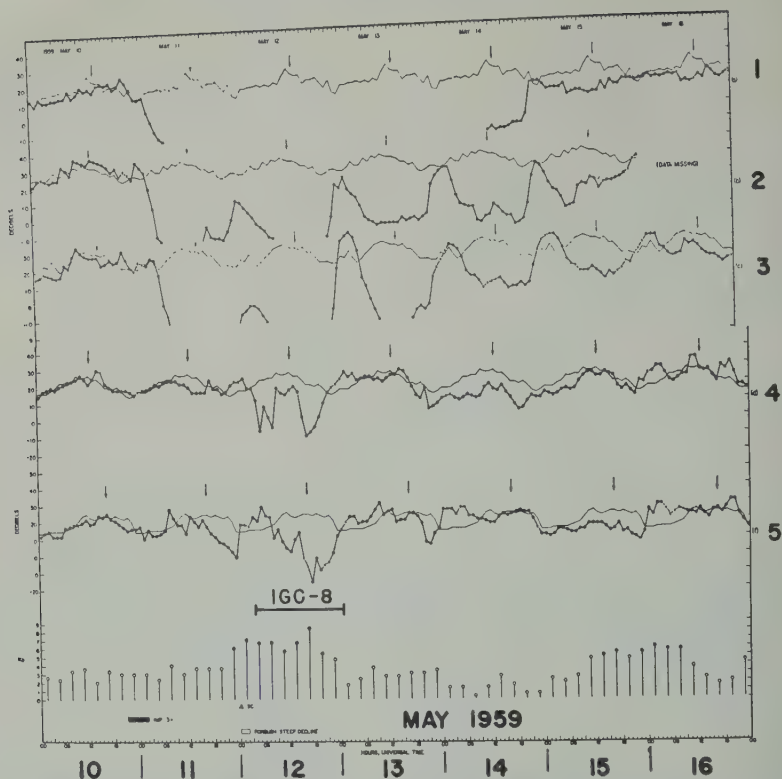


Fig. 10. Ionospheric absorption shown by the polar-cap scatter communications network signal (light line, normal signal; heavy line, signal during May 1959 at the indicated dates. From top to bottom, communication paths are as follows: (1) Thule to Sanderstrom; (2) Keflavik to Sanderstrom; (3) Goose Bay to Sanderstrom; (4) Grindavik to Kingston Wood, England; (5) Wachusett to Goose Bay. Note the great absorption over the high-latitude routes but the appreciable absorption at low latitudes only on May 12 at the time of the geomagnetic disturbances when cosmic-ray particles are seen at the latitude of Minneapolis. The time duration of flight IGC-8 on this scale is indicated by the horizontal bar. Courtesy of Dana Bailey.

detail in the legend, but they represent circuits over the polar region in the upper part of the figure, and in the lower part, circuits that go below the auroral zone. We see that, in the polar-cap circuits, the blackout began a few hours after the flare and persisted for more than 5 days. In the intermediate latitudes, the signal appears and reappears with a diurnal effect due to the detachment by sunlight of electrons from oxygen molecular ions. These ions are continuously produced by attachment of secondary electrons formed by the cosmic-ray particles. Below the auroral zone, in the bottom two curves, we see an effect only following the sudden commencement of the magnetic storm. The time of the balloon flight IGC-8 is shown by the horizontal black line just below curve 5 of Figure

10. It can be seen that this flight correlates with the time at which the solar protons were observed at low geomagnetic latitude, as shown by the partial absorption of the radio signals in circuits 4 and 5. This confirms the previously stated result that, at low latitudes like Minneapolis, the particles are allowed to enter during a disturbed condition of the geomagnetic field associated with a magnetic storm and the normal Störmer cutoffs are not effective at this latitude during that time.

We consider now the results obtained with the scintillation counter. This instrument gives information about the photon component of the storm at Minneapolis because of its high efficiency for photons compared with the other instruments flown. The scintillator, described

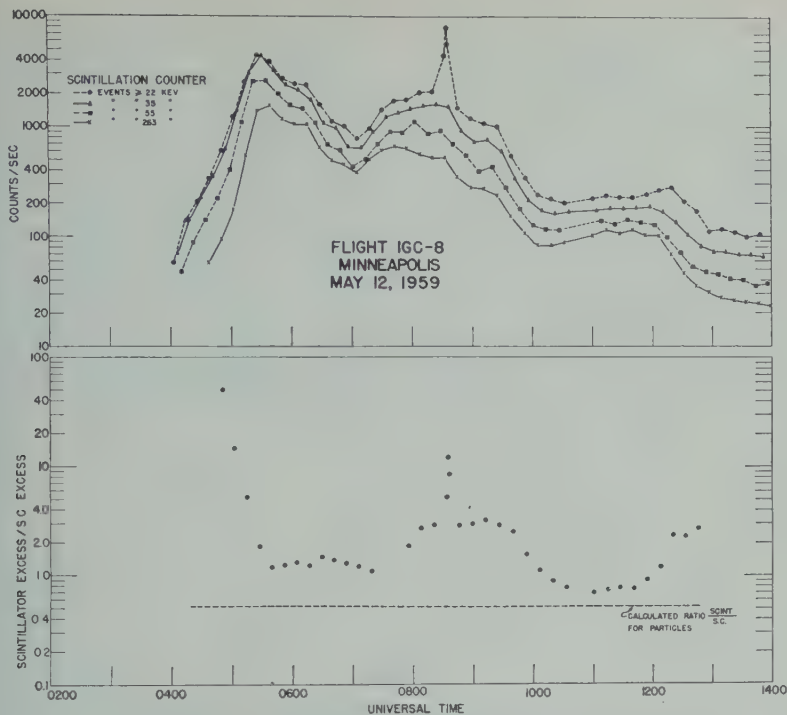


FIG. 1. Detailed analysis of the scintillation counter record. Upper section: counting rate of the scintillator for four different integral pulse-height energies located at 22, 35, 55, and 263 kev, uncorrected for particle effects. Lower section: ratio of scintillator excess to single-counter excess. This ratio shows photonic effects resulting only from the solar proton bombardment of the atmosphere and from X rays.

The instrument, in some detail, was equipped with an discriminator. The discrimination level was varied by a rotating potentiometer driven by a motor during the balloon flight so that in 10 minutes the instrument scanned from energies above 22 kev to energies above 263 kev. This covers the range of interest for such phenomena as auroral X rays. Of course, the scintillator also counts particles that have large energy loss in the scintillator less than 8 kev.

The dotted curve in Figure 4 shows the counting rate of the scintillation counter compared with other instruments. The dots represent the summing of each 10-minute scan, which includes all pulses greater than 22 kev in energy. Several features can be seen in the plot. First, the scintillation counter

showed a relatively larger effect at lower altitude than the counter or ion chamber and increased more slowly with increasing balloon height. When the balloon reached ceiling altitude, the scintillator tended to follow the other instruments but showed some additional effects, especially at about 0835 UT, when a very large peak occurred in which the counting rate increased to 8000 counts/sec coincident with a strong auroral maximum. All through the period between 0730 and 0930 UT, the scintillator showed evidence for an increased photon flux.

Before considering the details of the aurora itself, we show in Figure 11 a detailed analysis of the scintillation counter record. In the upper section of the figure is the counting rate of the scintillator for four different integral pulse-height energies located at 22, 35, 55, and 263 kev. These are the total rates uncorrected for normal cosmic-ray background or for the particle effects on this flight, so this is the summation

of effects due to all cosmic rays plus all photons of energy above that of the edge discriminator. The major effect, most of the time, is due to the large flux of the solar protons and the photonic effects they produce in the atmosphere, but considerable differences associated with the photonic component of the aurora appear in this record. The photonic component in the atmosphere is generated by the excess flux of protons as a result of nuclear collision and other processes. It seems to be generally of higher energy than the aurora. For example, on ascent of the balloon between 0400 and 0530 UT, there is very little photonic component in the interval between 22 and 35 kev, whereas later in the flight during the auroral maximum it develops strongly. (The large auroral peak at 0835 UT shows only on the 22-kev channel, but this is because the time variation of the peak is so sharp that, compared with the speed of scanning, the event came and went while the scanner was only in the neighborhood of the 22-kev position.) The scintillation rate returned toward the galactic cosmic-ray background by about 1400 UT, when the instrument ceased to give further information. The observation of the photon spectrum at various pulse-height intervals permits an evaluation of the spectrum all during the flight, but this analysis and other details about the source of the photons will be reserved for a future paper by one of us (P. D. B.).

In the lower part of Figure 11 is shown the ratio of the scintillator excess to the single counter excess. The excess is defined as the total observed rate on flight IGC-8 on May 12 minus the total observed rate on a quiet day, which in this case was IGC-7 on May 7, 1959, when it was known that no solar events were in progress. The subtraction of the quiet-day values removes particles and photons produced by the galactic cosmic rays in the atmosphere. The horizontal dotted line in the lower part of Figure 11 is the calculated ratio of scintillation counter to single counter for particles, which has a value very close to 0.5 because the scintillator has a geometrical factor almost exactly one-half that of the Geiger counter. The dots above this line represent 22-kev points once every 10 minutes throughout the usable portion of the data. It is seen that the ratio is in excess of that which can

be attributed to ionizing particles all through the flight, with several major increases, occurring during the ascent of the balloon. About 0540 UT, a broad maximum with a sharp peak at 0835 UT, and then a further increase after 1200 UT. For normal galactic cosmic rays the scintillator-to-single-counter ratio is about 3.0, which represent the extra response of the scintillator to photons generated by the primary cosmic rays in the atmosphere. We see that several times during the flight the ratio of excesses dips well below this value, implying that the solar protons on the whole generate more photons per particle than normal galactic cosmic rays per particle. This is probably reasonable view of the much lower energy of the solar protons. These particles lose most of their energy by ionization loss in the atmosphere rather than by nuclear processes, which would produce  $\pi^0$  mesons and photons of various energies. An observation—that at times the ratio dips below that which one would expect due to particles alone when the proton flux was known to be present with considerable intensity—implies that the high values of the scintillator excess to single-counter excess seen in the periods 0400 to 0530 UT, around 0800 to 1000 UT, and after 1200 UT, are correlated with the auroral phenomena.

Visual observations of the aurora were made throughout the period of the flight, and such correlation is indeed observed. It was noted before the launching of the balloon, beginning at 0330 UT, there was a very bright auroral rayed arcs and red color extending about halfway to the zenith as seen at Minneapolis. It faded at the time of balloon ascent. An auroral maximum developed beginning after 0600 UT and reached great intensity with strong green and red color observed at the zenith about 0835 UT. The aurora became sporadic around 1000 UT and receded to about 45° elevation in the north at about 0910 UT. No visual observations could be made during the radiation increase noted at 1200 UT because of the increasing twilight.

For a more exact comparison with the auroral features, we have examined the auroral all-sky camera records from stations adjacent to Minneapolis. (Films obtained from IGY Data Center for Aurorae, Fairbanks, Alaska.) In 1





Fig. 2. Auroral all-sky-camera records of the May 12 aurora from stations adjacent to Minneapolis. Upper two sections: early auroral buildup with maximum at 0535 UT. Lower two sections: great auroral buildup during balloon flight with maximum at 0836 UT. These photographs correlate well with visual observations at Minneapolis. The maxima agree within 1 minute with the time of the X-ray bursts accompanying the aurora.

shown the observations at Choteau, Montana, at the same geomagnetic latitude as Minneapolis but approximately 500 miles west, and a station at Fargo, North Dakota, close to Minneapolis. The upper sequence shows the aurora between 0500 and 0600 UT. At Choteau, the aurora appeared on the northern horizon, moved to its southern limit at approximately 0535 UT, and then receded northward or disappeared. The southern limit was north of the horizon at Choteau. The Fargo stations show similar behavior, but the aurora produced a bright feature over the sky near the zenith at 0529 UT. The Fargo pictures show some evidence of absorption, and distinct auroral features cannot be observed. The next auroral maximum occurred between 0800 and 0930 UT, with a very

great buildup close to 0835 UT as can be seen at both Choteau and Fargo; this is precisely the time of the sharp peak on the scintillation counter record. Such sharp maxima in the auroral X rays have been observed on numerous other occasions [Winckler, 1960]. The sharpness of the peak, which lasts only a few minutes, is considered to be due to the rapid motion toward and away from the zenith by a prominent feature of the aurora that is moving across the zenith with a velocity between 0.5 and 1 km/sec. This rate of motion combined with the rapid absorption of X rays in the atmosphere results in a short duration of the maximum as observed with the scintillation counter. We can consider this burst of X rays at 0835 UT to be a typical strong case. The scintillation counter,

which on this flight was flown for the first time during a strong aurora at Minneapolis, gave relatively huge counting rates, compared with the ion chamber and counter flown on other occasions, because of its greatly enhanced sensitivity to the X-ray quanta.

*Interpretation.* We now consider that the auroral X rays, the aurora itself, and the incidence of the solar protons at normally forbidden latitudes occur in a connected way somewhat as follows:

To begin with, the earth is immersed shortly after the time of the solar flare in a stream of the solar cosmic rays having energies from possibly 30 to 500 Mev and velocities averaging around one-third the speed of light. The energy density of these particles is very small, and is insufficient to perturb the magnetic field of the earth appreciably. These particles enter into the earth's geomagnetic field along Störmer orbits in the classical way and are allowed over a range of latitudes close to the pole. This behavior early in the event is clearly shown by Bailey's data (Fig. 10). The cloud of low-energy solar gas that is supposed to be the source of the geomagnetic storm arrives 1 day after the flare, has a very high energy density and a velocity of around 1000 km/sec, and is highly ionized. Upon the incidence of this gas on the earth's field, a serious perturbation occurs. Obviously, connections are established in some way between the internal parts of the field and infinity, or at least a connection is made between the gas stream from the sun and the internal parts of the earth's field. This connection permits the gas to descend into the atmosphere, producing certain features of the visible aurora. We assume that the gas may descend easily into the auroral zone, which connects with the outer reaches of the earth's field, but it may also penetrate into the field and descend at lower latitudes and produce the subauroral-zone visible aurora.

The subauroral-zone visible aurora is always observed to be accompanied by strong X-ray bursts, which we now believe may come from electrons discharged from the outer radiation belt of the earth by the perturbation of the solar gas and the ensuing fluctuations of the earth's field. Quantitative measurements of this effect have been made in the course of the history of satellite Explorer VI, in which it was

shown that at the time of a magnetic storm dumping of the electrons from the outer radiation belt occurred in sufficient number to produce the auroral X rays in a strong subauroral-zone aurora [Arnoldy, Hoffman, and Wine, 1960]. Such strong X rays are not seen in auroral-zone auroras [Anderson, 1960]. This is related with the fact that normally the intensity of the outer radiation belt is relatively low on a line of force of the magnetic field which connects with the atmosphere in the auroral-zone region.

The temporary connection established between the internal parts of the earth's field and the solar stream would naturally permit low-energy solar cosmic rays to enter into the same region as the solar plasma and associated aurora. It can readily be seen from an examination of Figure 4 and the auroral notes discussed above that there is a reasonably close correlation between the presence of the aurora in the auroral zone and the incidence of the low-energy cosmic rays below their normally allowed latitude.

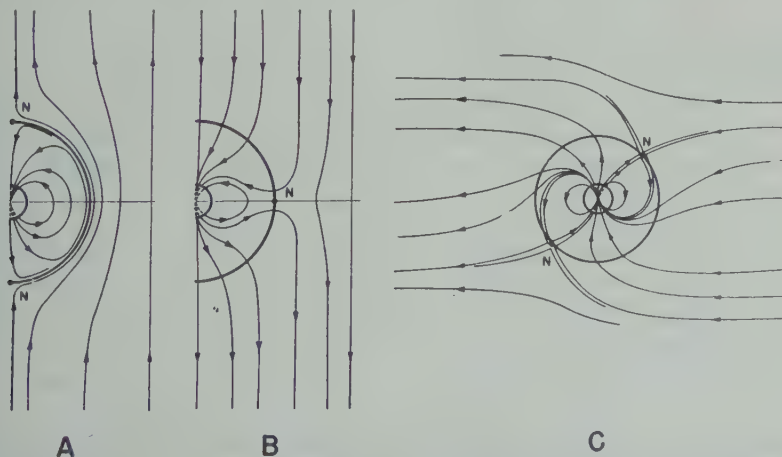
Several authors have contributed ideas as to the mechanism of changing the cosmic-ray cutoff energies. Rothwell [1959] has computed the minimum latitude at which the low-energy cosmic rays should be incident during a magnetic storm under three assumptions: (1) the magnetic storm cloud of solar gas can push the earth's field to within  $x$  earth radii of the dipole origin in the equatorial plane (the larger the storm, the smaller the value of  $x$ ); (2) particles can move into the earth's field at  $x$  earth radii via the storm cloud; (3) the earth's field is little affected by the storm cloud at distances less than  $x$  earth radii from the dipole center. Rothwell then uses Störmer theory to determine at what latitude protons of a given rigidity can reach the surface. On the assumption that at the latitude of Minneapolis the normal cutoff energy of primary cosmic-ray protons is 320 Mev, at the time of the magnetic storm which perturbs the earth's magnetic field to within  $x$  earth radii of the dipole center, protons of 320 Mev may reach the surface. Minneapolis is considered to be located at the geomagnetic latitude of  $55.4^\circ$  on the centered dipole approximation. However, directly measured cosmic-ray cutoffs increase this latitude to  $59^\circ$ , and if the lat

ed in terms of the magnetic dip, the ent latitude is  $61^\circ$ ; this latter latitude y by Rothwell in the calculations. Roth- theory approximately accounts for certain is of the observations. However, it is gen- sidered that the action of the solar cloud is to compress the lines of force of th's field in front of the advancing cloud t, although the cloud may reach to within radii of the dipole center, the Störmer applicable are not those characteristic erting particles at this distance. In Roth- theory it must be assumed that the cloud ates the field to the required distance of, arth radii, but that the lines of force are iously compressed, so that the earth's ot only undisturbed inside the 5-earth- egion but is on the whole not greatly l outside this region, either. It may be fficult to reconcile these two conditions, ng on the values of conductivity assumed ionized gas cloud incident on the earth's ifield.

ashi and Hakura [1960] have calculated ar type of perturbation of the Störmer but have carried out the analysis in more t. They suppose that a uniform magnetic resent around the earth which in the age is directed along the dipole axis to- eographic north and in the second along ole axis toward geographic south. The se corresponds to the initial part of a

magnetic storm in which the external terrestrial field is increased by compression or current systems from outside; the second, to the negative phase of the storm in which the earth's field is diluted. The mathematical result of superposing the uniform field on the dipole field is, in the first case, to reduce the total radial component to zero on a circle of radius  $R_0$ —which Obayashi and Hakura consider to be the limit of penetration of a solar plasma—which wipes away the field outside this point. The second case corresponds to the elimination of the  $\phi$  or latitude component of the magnetic field along a surface having again a certain radius  $R_0$ —which Obayashi and Hakura interpret physically as the distance of penetration of the solar plasma cloud. Störmer theory is modified by the addition of this uniform field, and the  $\gamma$  values (i.e., angular momentum constants) applicable are obtained and cutoffs computed. We reproduce here from the paper of Obayashi and Hakura the earth's field as it appears modified in the two cases in Figures 13A and B and in Figure 14, with the cutoff rigidities for incident particles expressed as a function of geomagnetic latitude during the negative phase of a geomagnetic storm. The parameters in this last set of curves refer to the radius of closest approach of the solar plasma cloud. This radius may be related to the value  $\Delta H_0$  of the storm field during the negative phase by the relation

$$\Delta H_0 = -M/R_0^3$$



3. Modification of the earth's dipole field by the addition of a uniform external field. A, external field parallel to dipole field; B, external field antiparallel to dipole field; C, external field perpendicular to dipole axis. Small circle shows size of the earth for a selected value of external field.



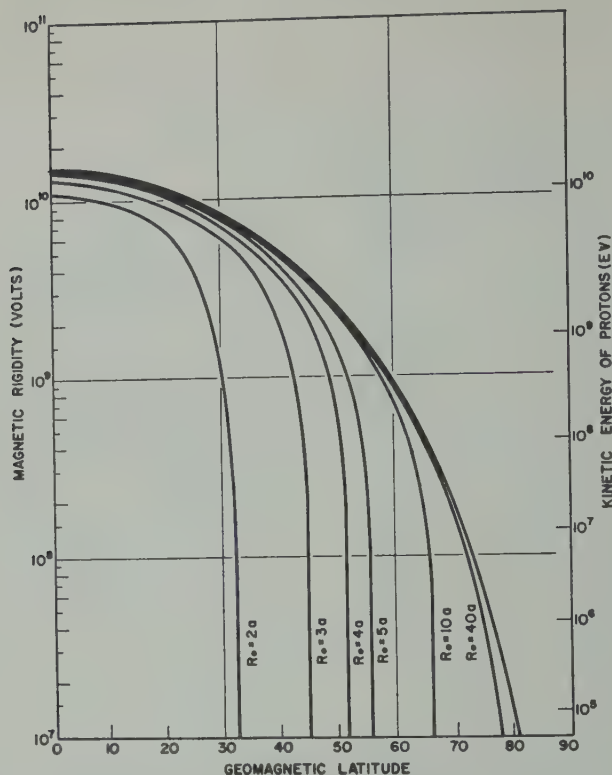


Fig. 14. Cutoff magnetic rigidity for the distorted geomagnetic field according to Obayashi and Hakura [1960].

Thus, in terms of the geomagnetic field changes, changes in cutoff may be inferred. We should like to emphasize here that Obayashi and Hakura have devised a means for relating observable quantities, which is a very useful procedure. No clear physical mechanism is suggested for the mathematical processes, however. Mathematically, one would get identical results with a uniform field extending to infinity instead of being contained inside the spherical surface as assumed by Obayashi and Hakura. The first case of Obayashi and Hakura actually is the closest to physical reality as it agrees with earlier notions about the inability of the earth's field to penetrate into the highly ionized interior of the solar cloud. The second case, however, in which the radial component continues out into the surrounding ionized medium, does not seem to have a definite physical basis. The radius of the sphere  $R_0$  is entirely a mathematical quantity, which can be eliminated completely from the theory by expressing everything in terms of

the imposed uniform field. The imposed uniform field represents the horizontal storm field at the equator. This field could be supplied, for example, from an external source such as a field or distant ring current without any plasma entering the situation.

The case of the modified cutoffs due to a ring current has been discussed by Ray [1960]. Ray computes in detail the alternation of Störmer theory by inserting a ring in the equatorial geomagnetic plane at a variable radius. He considers the exact field of the ring using elliptical functions, and finds that the usual allowed and forbidden regions as given by Störmer's theory are altered so that a second 'pass' occurs outside the ring. For a certain range of ring radii the 'outer pass' governs the admission of particles from infinity; for small ring radii the 'inner pass' governs it. The ring current is taken in such a direction as to reduce the cutoffs from their normal Störmer values at each latitude. The analysis is applied to primary cosmic rays.

eters the cutoffs in such a way as to approximate the observed 'knee' in the latitude dependence of cosmic-ray intensity.

For the fields close to the earth generated by the currents of various complexities, it is generally considered that, during the time of magnetic storms, magnetic fields of solar origin exist in the neighborhood of the earth. In the case of cosmic-ray accelerations by solar flares, for example, it is observed that the Forbush-type decreases of galactic cosmic rays occur simultaneously with the terrestrial geomagnetic storm. It is generally estimated that Forbush events may be caused by fields of 0.1 to 1  $\gamma$  extending over distances of the order of 1 astronomical unit. If these fields had magnitudes of 100  $\gamma$ , the observed cutoffs in cosmic-ray cutoffs during magnetic storms could be produced. With this in mind, for comparison, in Figure 13C, the case of such a field exists perpendicular to the ecliptic axis. Here an asymmetry is produced between the northern and southern hemispheres. Neutral points are located midway between the equator and the polar axis. Again in this case, in the others, an external field of 100  $\gamma$  would permit the solar cosmic rays to be incident at Minneapolis.

The production of such very strong solar fields of  $\gamma$ -ray strength near the earth seems very unlikely. Trapped plasma fields of this strength blown out from the sun may be possible, however. These problems will best be settled by direct magnetic measurements with satellites and space probes in the future.

**Acknowledgments.** We are indebted to Harold Huch, of the Geophysical Institute, University of Minnesota, for alerting our group that a cosmic-ray study was in progress. L. Peterson, W. Huch, T. Peterson, and many student members of the balloon group assisted with the preparation and launching of the balloon flight.

This work was assisted by the joint program of the Office of Naval Research and the U. S. Atomic Energy Commission under contract Nonr-710(19).

#### REFERENCES

Arnoldy, K. A., Balloon observations of X rays from the auroral zone I, *J. Geophys. Research*, **65**, 1364, 1960.  
 Arnoldy, K. A., R. Arnoldy, R. Hoffman, L. Peterson, and J. R. Winckler, Observations of low-energy solar cosmic rays from the flare of

22 August 1958, *J. Geophys. Research*, **64**, 1133-1147, 1959.  
 Arnoldy, R. L., R. A. Hoffman, and J. R. Winckler, Observations of the Van Allen radiation regions during August and September 1959, I, *J. Geophys. Research*, **65**, 1361-1376, 1960.  
 Charakhchian, A. N., V. F. Tulinov, and T. N. Charakhchian, A large disturbance in the intensity of cosmic rays in the stratosphere 11-16 May 1959, *J. Exptl. Theoret. Phys. USSR*, in press 1960.  
 Denisse, J. F., Relation entre les émissions de rayons cosmiques solaires et certains sursauts radioélectriques, *Paris Symposium on Radio Astronomy*, edited by R. N. Bracewell, Stanford University Press, pp. 237-239, 1959.  
 Denisse, J. F., Certain aspects of solar activity in relation to cosmic ray production by the sun, *Proc. First Intern. Conf. on Space Sci., Nice, January 1960*, in press, 1960.  
 Erickson, W. C., Radio emission in the outer corona, *Phys. Rev. Letters*, **3**, 365-367, 1959.  
 Fowler, P., P. S. Freier, and E. P. Ney, The primary alpha particle spectrum over North America and geomagnetic cut-off energies, *Nuovo cimento, Suppl.*, **8**, 492-499, 1958.  
 McDonald, F. B., Study of geomagnetic cut-off energies and temporal variation of the primary cosmic radiation, *Phys. Rev.*, **107**, 1386-1395, 1957.  
 McDonald, F. B., and W. R. Webber, Proton component of the primary cosmic radiation, *Phys. Rev.*, **115**, 194-205, 1959.  
 Ney, E. P., J. R. Winckler, and P. S. Freier, Protons from the sun on May 12, 1959, *Phys. Rev. Letters*, **3**, 183-185, 1959.  
 Obayashi, T., and Y. Hakura, Enhanced ionization in the polar ionosphere caused by solar corpuscular emissions, *J. Radio Research Lab.*, **7**, 27-66, 1960.  
 Ray, E. C., Effects of a ring current on cosmic radiation, *Memoria del V Congreso Internacional de Radiación Cósmica, Guanajuato, 5-13 Sept., 1955*, Instituto Nacional de Investigación Científica, Mexico, pp. 106-121, 565 pp., 1958.  
 Reid, G. C., and H. Leinbach, Low-energy cosmic ray events associated with solar flares, *J. Geophys. Research*, **64**, 1801-1805, 1959.  
 Rothwell, P., Magnetic cutoff rigidities of charged particles in the earth's field at times of magnetic storms, *J. Geophys. Research*, **64**, 2026-2029, 1959.  
 Winckler, J. R., Balloon study of high-altitude radiations during the IGY, *J. Geophys. Research*, **65**, 1331-1359, 1960.  
 Winckler, J. R., L. Peterson, R. Arnoldy, and R. Hoffman, X rays from visible aurorae at Minneapolis, *Phys. Rev.*, **110**, 1221-1231, 1958.  
 Winckler, J. R., L. E. Peterson, and R. L. Howard, Balloon gear monitors cosmic radiation, *Electronics*, **31**, 76-79, 1958.

(Manuscript received June 23, 1960.)





## Observations of Solar Cosmic Rays Near the North Magnetic Pole

K. A. ANDERSON AND D. C. ENEMARK

*Department of Physics and Astronomy  
State University of Iowa  
Iowa City, Iowa*

**Abstract.** Following a large solar flare on July 16, 1959, fluxes of protons in the energy region 85 to 300 Mev in a steep spectral distribution were observed by means of high-altitude balloons over Resolute Bay, Canada. The flare particle intensity decreased rapidly and quite steadily over a period of several days, but on the occasion of the last flight, launched on July 27, a small flux remained above the high-energy cosmic-ray level. At that time the active region in which the flare occurred was nearly opposite central meridian. The time dependence of the particle intensity followed quite well a  $t^{-2}$  law for which  $t$  equaled zero about 16 hours after the flare. This low-energy cosmic-ray-producing flare was accompanied by large and complex radio noise emission, and the particles produced the usual terrestrial effects of long-enduring polar-cap radio blackout, including large riometer effects. During the magnetic storm, at the time the particles were being admitted at forbidden geomagnetic latitudes to the south, the particle intensity was undisturbed over Resolute Bay.

**Introduction.** During July 1959 several high-altitude balloons carrying cosmic-ray apparatus were launched from Resolute Bay, Northwest Territories, Canada, which is at 74.7°N geographic latitude and 94.9°W longitude. The North magnetic dip pole is about 150 km to the southwest of Resolute Bay. The first flight, launched late on July 17, encountered particle fluxes, identified as predominantly composed of protons that first became detectable at an atmospheric depth of about 100 g cm<sup>-2</sup>, then increased rapidly in intensity as the balloon ascended, until at 6 g cm<sup>-2</sup> the flux was several hundred times greater than the normal galactic cosmic-ray level. Balloon launchings were continued until July 27; on each flight the particles were found to be still present but steadily decreasing in intensity. Since the origin of the particles is believed to be associated with the large solar flare that appeared late on July 16, we are dealing with one of the low-energy solar cosmic-ray events concerning which much information has recently been acquired [Anderson, 1958b; Freier, Ney, and Winckler, 1959; Anderson, Arnoldy, Hoffman, Peterson, and Winckler, 1959; Reid and Leinbach, 1959; Ney, Winckler, and Freier, 1959; Brown and D'Arcy, 1960].

The sequence of solar and terrestrial events.

The large chromospheric flare that occurred on July 16 began at 2115 UT and was located at N16 W29 heliographic coordinates, according to the Central Radio Propagation Laboratory monthly tabulation of solar-geophysical data. The peak optical development was attained at 2129 UT, and about 3 hours later the flare ended. This flare produced great amounts of radio-frequency energy. Following fast drift bursts from 2120 to 2122 UT there appeared continuum radiation confined mainly to higher frequencies, above about 150 Mc/s (Maxwell, private communication). The duration of the continuum emission was at least 4.5 hours at 167 Mc/s. At 2800 Mc/s the noise outburst was of very great intensity, the peak received flux reaching  $6500 \times 10^{-22}$  watt m<sup>-2</sup> c<sup>-1</sup> sec<sup>-1</sup> at 2154 UT. At 18 Mc/s, the lowest radio frequency for which observations are available, very intense bursts were present from 2122 to 2155 UT. Their peak energy fluxes were at least as high as  $10^{-18}$  watt m<sup>-2</sup> c<sup>-1</sup> sec<sup>-1</sup> (Warwick, private communication). Beginning at 2200 UT the solar emission at 18 Mc/s became somewhat smoother; presumably it is part of the continuum spectrum with its appearance delayed about 40 minutes with respect to the higher frequencies. This radiation was rather weak and could be detected only for about 1 hour.

The arrival at the earth of the nucleonic particles occurred within  $1\frac{1}{2}$  hours of the flare peak [Reid and Leinbach, 1959]. The principal terrestrial effect produced by the solar proton flux was enhanced ionization in the ionosphere, which, owing to the penetrating nature of the particles, reaches the *D* layer and below, where the electron collisional frequencies are high. This effect manifests itself through interruption of VHF radio communication paths and marked attenuation of cosmic radio noise over the polar regions. The disappearance of these effects somewhat south of the auroral zone is due to the rather low energies of the protons, so that the earth's magnetic field cuts them off south of those latitudes; the abruptness of the effect is due to the very steep energy spectrum of the solar particles.

The effect on radio communications during this particular event began, according to Bailey (private communication), early on July 17 and continued until July 20. The maximum riometer effect (Leinbach, private communication) due to these solar particles was greater than 15 db, and enhanced absorption of the 28-Mc/s cosmic noise could be detected until about 1200 UT on July 19 at College, Alaska, somewhat to the south of the auroral zone.

The sudden-commencement magnetic storm associated with the large flare on July 16 began only 19 hours later, at 1638 UT on July 17. It lasted for about 10 hours and was responsible for producing a Forbush decrease that began about 1700 UT and decreased the rate in the sea-level neutron monitor at Resolute Bay by about 16 per cent [Wilson, Rose, and Pomerantz, 1960].

Before the event just described, two large solar flares that occurred on July 10 and 14 also produced low-energy cosmic rays, which were extensively studied by means of balloon-borne equipment over Minneapolis, Fairbanks, Fort Churchill, Chicago, Germany, and Russia [Charakhch'yan, Tulinov, and Charakhch'yan, 1960] by numerous investigators. The occurrence of several solar proton emissions within the period of a few days had previously been noted; three [Rothwell and McIlwain, 1959] or possibly four [Reid and Leinbach, 1959] low-energy cosmic-ray events occurred between August 16 and 26, 1958.

*Time variations of the solar proton flux.* Fig-

ure 1 shows the counting rate of the single Geiger counter during the first flight at Resolute Bay, which was launched about 25 hours after the occurrence of the flare. More than 29 hours of data were obtained at an atmospheric depth of  $6 \text{ g cm}^{-2}$ . The depth of the balloon changed by no more than  $0.3 \text{ g cm}^{-2}$  during the course of the flight, so that the counting rate shown there gives a quite accurate picture of the real time variations in the particle flux. The dead-time correction, which amounts to about 20 per cent at the highest rates, has not been made. Some of the main features of interest during the flight are:

1. The particle flux decreases in a general steady manner, although at times, for periods of a few hours, the counting rate may remain more or less constant or even increase.
2. The particle flux is unaffected by variations in the magnetic field during the latter part of the sudden-commencement storm and also during the subsequent geomagnetic disturbances.
3. When the protons are let in at southern forbidden latitudes by the magnetic storm effect the flux of protons above Resolute Bay is undisturbed.

The riometer record taken at King Salmon (Leinbach, private communication), several degrees below the auroral zone in southern Alaska, indicates the appearance of the protons by detection of their ionization in the atmosphere. In Figure 1 the protons are seen to appear between 0300 and 0700 UT on July 18.

Figure 2 shows the response of all three detectors over the first part of the balloon flight made on July 21-22. This flight continues to show the same general decrease of the flare particle flux but again with occasional periods where the intensity changes by very little.

Figure 3 shows the presence of the flare particles during July 24-25. Their intensity is not only about 15 per cent of the galactic cosmic-ray background, but they can be easily recognized by the characteristic upturn in the ion-chamber pulse rate at the smallest atmospheric depths attained by the balloon during its ascent. For this particular flight no effects other than the normal cosmic-ray background became apparent until a depth of somewhat less than  $1 \text{ g cm}^{-2}$  had been reached. The ion-chamber rate then continued to increase until the ceiling at

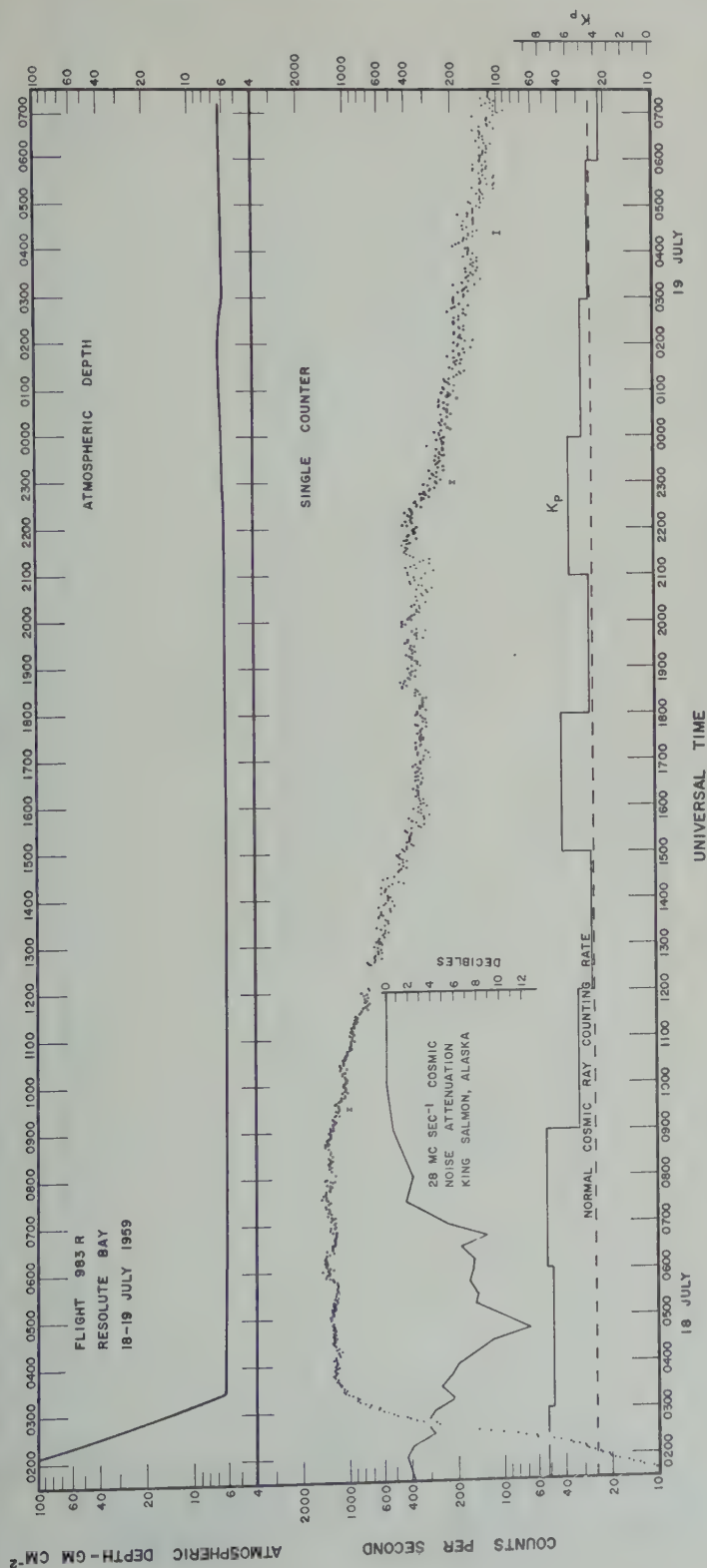


Fig. 1. The single counter rate during the first flight at Resolute Bay. Also shown is the pressure depth maintained by the balloon. The riometer record from King Salmon, Alaska (courtesy of H. Leinbach), shows the letting in of these particles at geomagnetically forbidden latitudes, an effect closely associated with high  $K_p$  values.



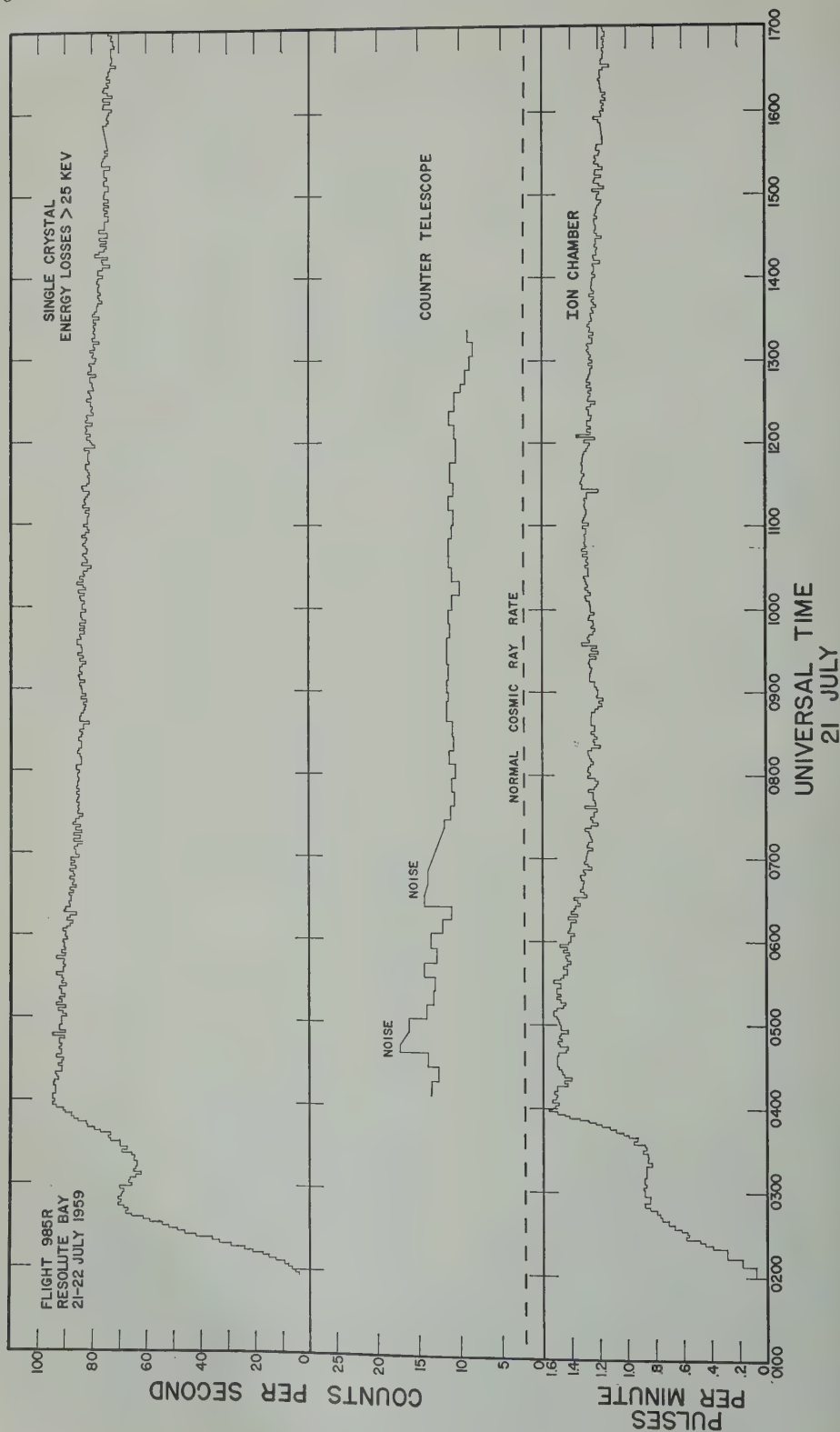
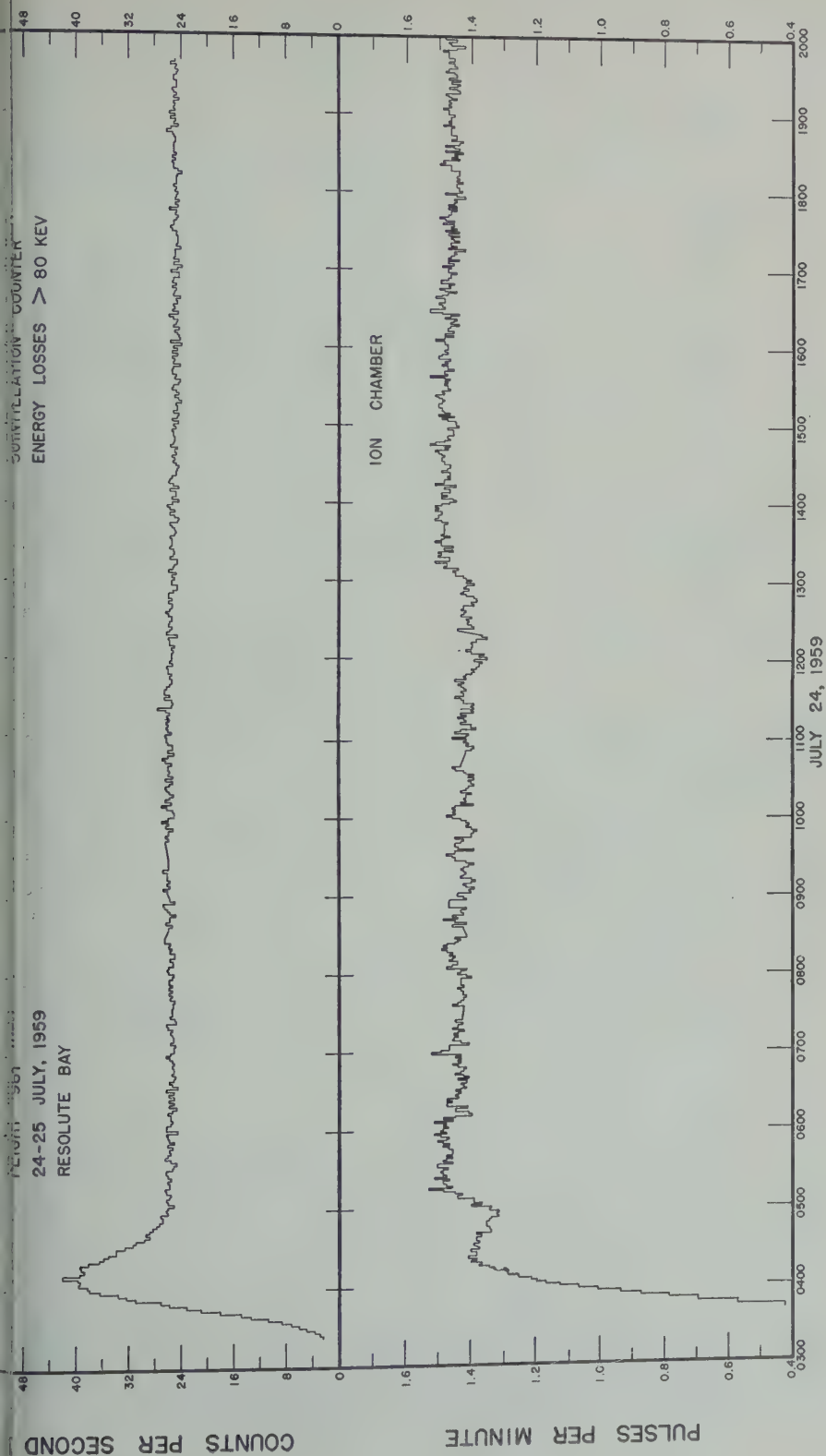


Fig. 2. Detector rates during the flight which began more than 4 days after the flare.



## UNIVERSAL TIME

Fig. 3. Seven days after the flare the flare particles can still be recognized by the characteristic upturn in the ion-chamber record. The abrupt change in the flux just past 1200 UT shows that in addition to the over-all  $t^{-3}$  decay some other control is exercised on the particles.

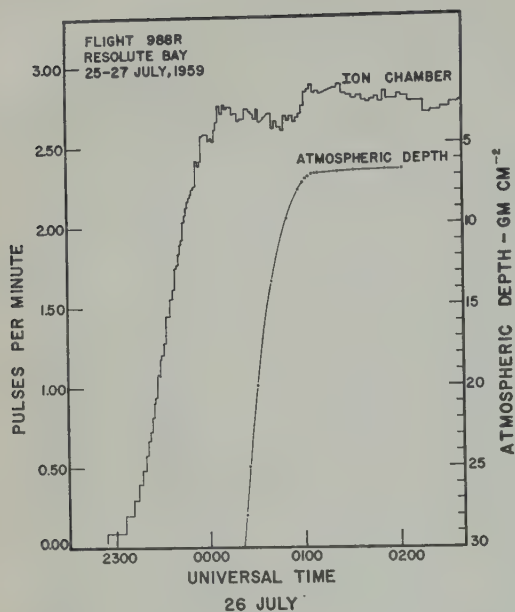


Fig. 4. This flight begins about 10 days after the flare. Forty hours later a small flux of particles can still be seen.

tude corresponding to  $7 \text{ g cm}^{-2}$  was attained. From range-energy considerations this means that protons having energy at least as high as 115 Mev and down to 90 Mev are still arriving at the earth nearly 9 days after the occurrence of the solar flare. At these late times the intensity of the particles has become so small that it is difficult to estimate by means of the single counter and counter telescope, because of inadequate statistical accuracy and their large transition effect. For the last two or three flights, therefore, it is not possible to reliably identify the particles as protons. Also, the steepness of the particle number-energy spectrum, in addition to the small particle flux at these times, does not allow the presence of the spectrum above 115 Mev to be verified, although there is no reason to believe that it is not there.

Figure 4 shows the first 4 hours, including the ascent, of the last flight made at Resolute Bay on July 27, more than 10 days after the solar flare. A small effect due to the flare particles can still be seen. The remainder of this flight is shown in Figure 5, which presents the ion-chamber data for all the Resolute flights. Plotted there is the excess over the ion-chamber rate due to the high-energy cosmic-ray back-

ground determined as accurately as possible each flight. Some of the associated solar terrestrial phenomena are also represented. The dashed curve is a  $t^{-3}$  law for which  $t$  equals zero about 16 hours after the beginning of flare. This curve fits the data in a general fashion over a 10-day period and provides a much better fit than  $t^{-2}$ ,  $t^{-4}$ , or  $e^{-at}$  laws would.

**Particle composition and energy spectrum**  
The instrumentation for each flight included an omnidirectional particle counter (either a Geiger tube or a single scintillation crystal), a vertically directed counter telescope, and an ionization chamber. An olland cycle device determined the pressure depth of the balloons every 2 minutes during the entire flight. By studying altitude as well as relative counting rates from the three particle detectors, we conclude that the solar flare particles observed over Resolute Bay beginning late on July 17, 1959, consisted of at least 85 per cent protons, with the measurements being consistent with a 100 per cent proton composition. The details of this method of particle identification have already been described [Anderson, 1958b; Anderson, Arnold, Hoffman, Peterson, and Winckler, 1959] and need not be repeated here. The above conclusion applies strictly for the first few days following the flare since after that time the particle flux becomes so small that the method is no longer very accurate.

The energy spectrum of the protons is found from the curve of ionization in the argon-filled ion chamber versus atmospheric depth by the method used previously in connection with the August 22, 1958, solar cosmic-ray event. The result of this analysis can be represented by the following expression for the differential kinetic energy spectrum

$$n(E) dE = E^{-4.5} dE \quad 85 < E < 400 \text{ Mev}$$

This result applied from 0200 to 0400 UT on July 18, approximately 28 hours after the flare. Corrections for energy loss in the atmosphere have been made so that all spectra given here apply at zero atmospheric depth. Owing to the very high flux of particles, secondary effects such as  $\gamma$ -ray and neutron production may be important. In this event, the spectrum found by this method would be in error, the exponent being larger than the value found above. The other determination of the energy spectrum



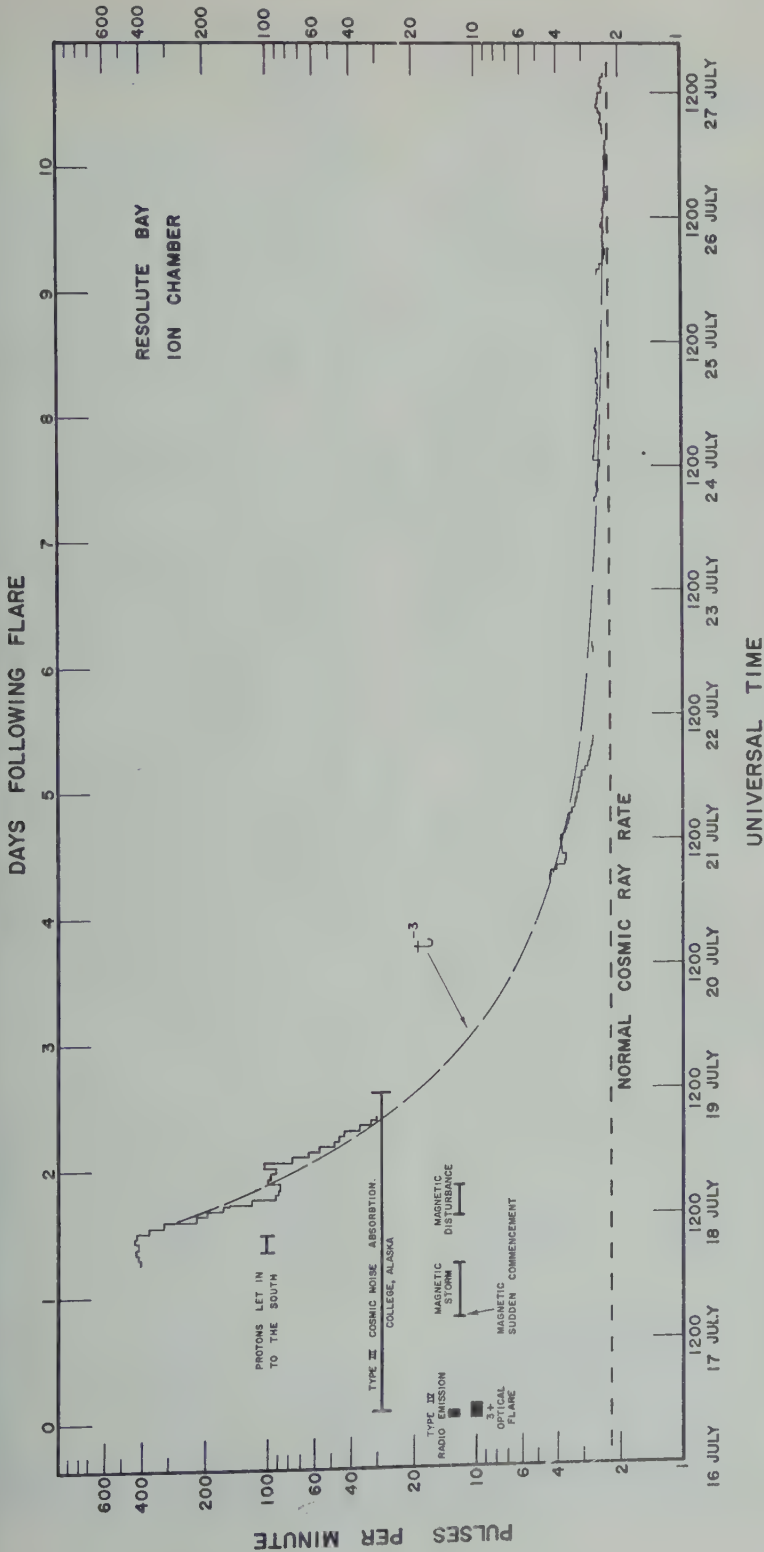


Fig. 5. Summary of the ion-chamber records obtained during the Resolute flights. These pertain to an atmospheric depth of  $6.5 \text{ g cm}^{-2}$ .

has been made by this method from 0310 UT to 0400 UT on July 21 at a time when the particle flux was quiescent and not very intense. On this occasion the energy spectrum can be represented by

$$n(E) dE = KE^{-6.1} dE \quad 85 < E < 300 \text{ Mev}$$

Neither the abundance nor the upper limits to the abundance of energetic protons and neutrons produced by the July 16 flare can be obtained from this event since no balloons were at high altitude during its optical development and, therefore, presumably at the time the particles were accelerated. As has already been reported, it was established that the low-energy cosmic-ray-producing flare of August 22, 1958, yielded undetectable quantities of these neutral radiations. The upper limits to the abundance above 100 kev and neutrons above 40 Mev were only a few per cent of the proton flux above 100 Mev or a few hundredths of a per cent of the protons above 30 Mev. The acceleration of the protons was thus electromagnetically clean at high energies, and furthermore nuclear or thermonuclear processes were not important for the production of the energetic protons reaching the earth.

Flight 985, made July 21-22, included a vertical counter telescope whose two coincidence elements were thin, circular plastic phosphors. This detector will be referred to as the energy-loss telescope. Its operation is as follows: When an energy loss greater than about one-third of the most probable relativistic  $\mu$ -meson energy loss occurs in both crystals within the resolving time of the circuits (about 1 microsecond), the output from the photomultiplier viewing the top crystal is passed on to a height-time converter and telemetered to the ground station [Enemark, 1959]. The range of this instrument was adjusted so that the  $Z = 1$  relativistic peak could be fully seen while the  $Z = 2$  peak would lie roughly in the mid-part of the energy-loss region. Thus the dynamic range was about 25, and energy losses up to about 8 times minimum occurred on a linear scale. The total counting rate of the telescope above the triggering level could readily be obtained, and for flight 985 this is plotted in Figure 2. Because of its good angular resolution the ratio of the counter-telescope rate to ion-chamber rate is about 5 times higher for the solar protons than for the galactic cosmic

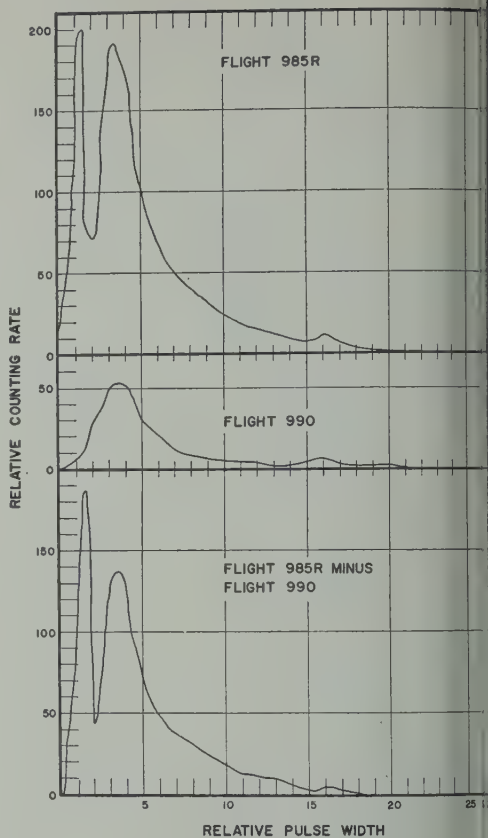


Fig. 6. The middle diagram shows the resolution obtainable with the energy-loss telescope when solar effects are absent and only the galactic cosmic-ray flux is present. When this is subtracted from the distribution during the appearance of the flare particles (top diagram) it is seen that fast helium is present.

rays. This is due to the strong collimation imposed on the more absorbable flare particles.

That the energy-loss telescope gives good resolution can be seen from Figure 6. The middle part of this figure shows a pulse height distribution obtained for the energetic galactic cosmic radiation when solar effects were entirely absent. The peak due to fast helium can be seen clearly. In the top part of the figure is shown the pulse height distribution obtained during flight 985. In this case about 90 per cent of the counts are due to the solar flare particles. When the normal cosmic-ray distribution is subtracted from that of flight 985, few if any pulses remain in the fast  $\alpha$ -particle region, as can be seen in

bottom part of Figure 6. During a considerable part of this flight, however, there was a  $\alpha$  particle in the fast  $\alpha$ -particle region which did not react away with the galactic cosmic-ray content. This we attribute to an effect of the electronic circuits which serves to shorten the range of the detector greatly. We believe that the best way to interpret the counter-poise results is to say that no fast  $\alpha$  particles were being detected in the flare radiation. But, since we cannot completely determine the nature of the electronic irregularity, the  $\alpha$ -particle results of the Minnesota group [Freier, 1957] are much more conclusive.

Considering that the protons we observe at 6  $\times 10^9$  cm depth must have had velocity greater than  $1.25 \times 10^{10}$  cm sec<sup>-1</sup> ( $\beta = 0.42$ ) before entering the atmosphere, we conclude that there must have been less than 1  $\alpha$  particle for each proton above this velocity.

The lowest-rigidity  $\alpha$  particle that has sufficient energy to penetrate the detector is 0.85 bv. The energy of a proton corresponding to this rigidity is 330 Mev. The proton energy spectrum is so steep that very few of the particles observed on flight 985 can have this energy. On the basis of the  $E^{-4}$  integral spectrum, fewer than 0.25 per cent of the particles have energy greater than 380 Mev, and so the best limit we can obtain on the basis of rigidity is that there must be less than 1  $\alpha$  particle for every 4 protons above the rigidity of 0.85 bv.

Using photographic emulsions Freier [1957] has identified helium in the flare particle flux, but its abundance at rigidities greater than about 0.5 bv is 1 for each 18 protons.

**Discussion.** 1. The identification of these proton fluxes over the magnetic pole of the earth cannot be taken as conclusive evidence that the majority of most of these particles cannot be connected with the geomagnetically trapped radiation. Many other arguments have pointed to this conclusion previously. They were based on the absence of auroral phenomena associated with the proton flux, their markedly different energy spectrum, detailed association with solar flare phenomena, and the presence of polar-cap ripple effects.

However, consideration should be given to the possibility whether the particles seen many days after the flare can be temporarily trapped in

the region of the earth. A special configuration of the magnetic field near the north pole would be required, but this, though highly unlikely, cannot be logically excluded. The fact that one mathematical expression, the  $t^{-3}$  law, fits the particle flux very well over the 10-day period suggests strongly that the protons have but one source and that geomagnetic trapping plays no role whatsoever.

2. The differential energy spectrum obtained for this event is very similar to the one found on August 22-23, 1958, and to spectra found subsequently for other events by various experimenters. In the proton energy region from about 85 to 400 Mev a power-law representation with an exponent between 4.5 and 6 is quite satisfactory. The multiplicative constant,  $K$ , however, is quite variable in these events. The spectrum derived for the high-energy solar cosmic-ray emission of February 23, 1956, has a similar exponent but applies to a range of energy, two orders of magnitude greater, or, roughly, 1 to 50 bev. Interestingly, there seems to be no occasion on which the sun produced this characteristic steep spectrum at intermediate energies. There is strong evidence, however, that the sun may at times emit a spectrum which lies at very low energies and contains an undetectable number of protons having energy greater than 100 Mev. Figure 7 shows an ionization-chamber record obtained on a balloon flight above Fort Churchill at an atmospheric depth of 10 g cm<sup>-2</sup> on August 29-30, 1957. During this flight marked polar-cap (type III) cosmic-noise absorption was present in the riometer at Fort Churchill. This has also been reported as a typical low-energy solar cosmic-ray event by Reid and Leinbach [1959], but on that occasion, the solar proton flux at the altitude of the balloon could not have been more than about 1 per cent of the normal cosmic-ray flux. This means that the particles producing the long-enduring cosmic-noise absorption must have been below 100 Mev in kinetic energy, if they were protons. The sporadic counting-rate effects seen on that flight have been positively identified as being due to X rays that appeared in connection with the geomagnetic storm on that occasion [Anderson, 1958a].

3. As has been noted the radio-frequency continuum emission associated with the July 16,



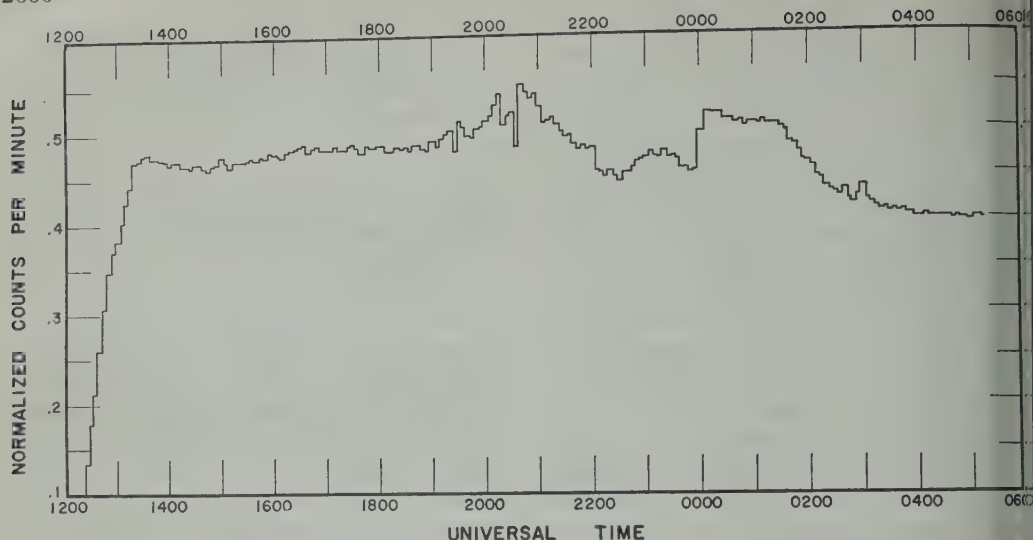


Fig. 7. Flight 61, made on August 29-30, 1957, shows that no solar particles were present at a depth of  $10\text{g cm}^{-2}$  even though large polar-cap riometer effects are observed at these times. The sporadic counting rate effects are due to soft X rays associated with the geomagnetic storm on that occasion.

1959, flare was cast toward higher frequencies than the August 22, 1958, cosmic-ray-producing flare. Since the rf continuum is believed to be closely associated with the proton flux [Anderson, 1958b; Anderson, Arnoldy, Hoffman, Peterson, and Winckler, 1959], it is important to discuss this situation. In particular there is the possibility that the protons and the electrons that produce the rf continuum by means of synchrotron radiation are accelerated in the same process at the same time. But the form of the proton spectra produced in these events is remarkably constant whereas the frequency at which the maximum continuum radiation takes place is variable. This behavior might be accounted for if the electrons are contained in magnetic fields varying in intensity from one event to another, thus altering the gyrofrequency. Electrons of the same energy in magnetic fields of differing strength would be expected to have different lifetimes, assuming that ionization loss is not important. The duration of the rf continuum seems to be much longer for the August 22, 1958, event when the gyrofrequency, and presumably the containing magnetic fields, were much lower, compared with the July 16, 1959, occurrence.

Another alternative that would account for

different rf features from solar flares which produce cosmic rays having similar spectra is the following: the protons are produced with spectra varying in form from one event to another, but the emission process from the solar atmosphere alters the spectra to the one observed at the earth. This does not appear to be a very likely possibility, however.

4. As late as 10.5 days following the large flare on July 16, protons with energy 90 to 1 Mev are still arriving over Resolute Bay. Since the spectrum is very steep and particle intensities small at this time, the presence of the higher-energy part of the spectrum cannot be directly verified. Also, particle identification is uncertain at these late times, but the observed flux is still taken to be composed predominantly of protons.

During this 10.5-day period following the flare, the particle flux has generally decreased in a regular fashion, although there are periods of the order of a few hours when the flux becomes relatively constant or increases somewhat. These are real effects and cannot be attributed to altitude changes of the balloons. It should be noted that the balloons were drifting westward at a rate of 8 to 12 knots. At the time during the Resolute observations did t

cosmic-ray flux reach zero, although it nearly did so on the last flight.

The over-all behavior of the particle flux is described by a  $t^{-3}$  law with  $t$  equaling zero hours after the flare peak. That means that the proton flux changed by a factor of 2 between the fourth and fifth days after the flare, at times later than this the rate of decrease is lower and at times earlier it is greater. Combining this time dependence with the spectrum found above, the following expression for the integral particle flux may be written:

$E, t)$

$$= 15 \times 10^{10} t^{-3} E^{-4} \text{ particles}/(\text{cm}^2 \text{ sec})$$

$$85 < E < 300 \text{ Mev} \quad t > 1.2 \text{ days}$$

where  $E$  is the proton kinetic energy in Mev and  $t$  is the time following the flare peak in days (2200 July 16). The exponent of the energy spectrum seems to be constant in time, but substantial changes would not be observed late in the event.

The departure from a smooth decline, particularly the periods when the flux increases, means that even several days after the flare some sort of control is being exercised over the particles. The fluctuations cannot be related to magnetic disturbances as detected by ground-level magnetometers, and they do not occur as periodic diurnal effects, so that the possibility of their being impact-zone effects is apparently ruled out.

The result that the geomagnetic storm associated with the solar flare has little if any effect on the low-energy solar particle flux (these particles merely continue their steady decline during this time) shows that no great quantity of low-energy protons can be trapped in the storm. Furthermore, the origin of the flare protons, which appear at forbidden geomagnetic latitudes for a few hours during the magnetic storm, is therefore not due to the release of particles from the solar gas cloud.

The Resolute results also establish that particles as low as 100 Mev in energy continue to arrive at the earth at times when very high-energy galactic cosmic rays are being removed and while these rays are recovering in intensity. These facts are of importance to theories of

the Forbush decrease, as is the observation that, in the August 22, 1958, solar cosmic-ray event, protons of 100 Mev energy and probably those of much less energy as well arrived at the earth in times roughly consistent with their speed. The rapid build-up of intensity of these particles in the August 22 event means that any ordered solar and interplanetary magnetic fields must be weak and any disordered magnetic fields present must permit very rapid diffusion of these particles.

6. Four and one-half days after the flare, the active region in which it appeared arrived at the solar limb, on the basis of a synodic rotation period at the latitude of this sunspot group of 27 days. A flight aloft during the passage of this active region behind the limb observed no unusual departure from the steady decline of the proton flux. Ten and one-half days after the flare a small but distinct flux of flare particles remained, and at this time the position of the active region was about  $10^\circ$  from passing central meridian but on the back side of the sun (heliographic longitude W170).

This behavior bears on the question of the storage and emission region of the protons. Arguments given previously [Anderson, Arnoldy, Hoffman, Peterson, and Winckler, 1959], suggested that, because of the close connection of the protons with the radio noise outburst that occurred for the August 22, 1958, event, the protons were being emitted over a period of days from the same region that contained energetic electrons which were the source of the rf continuum via synchrotron radiation. On this occasion, measurements by Denisse indicated that the radio region was very large and located high in the solar atmosphere.

This view about the emission of the flare particles has also been taken by Erickson [1959] in connection with the May 12, 1959, low-energy solar cosmic-ray emission. His measurements on this occasion showed long-lived and large radio-noise sources which apparently rotated with the sun. Rotation of these regions with the sun would not be inconsistent with the observation that the proton flux does not disappear after the active region reaches the solar limb, since only a few per cent of the source area would be occulted by the sun, and this rather slowly. We would, therefore, not expect

any marked effect on the particle flux if it was coming from this region. On the other hand, this view implies that when a low-energy cosmic-ray-producing flare occurs on the back side of the sun some of this particle flux must reach the earth.

Bailey has recently reported (private communication) that a few polar-cap VHF radio wave absorption events have occurred without chromospheric flare or sizable radio-noise emission. From this, and from their gradual onset, he believes that they can best be interpreted as effects from back-side solar flares. Such an interpretation further supports our view that the flare particles are temporarily trapped in the solar atmosphere by magnetic regions having a volume of the order of the sun itself or several times larger. As previously discussed, these regions have been identified with the long-lived sources of continuum RF emission.

7. Reid and Leinbach [1959] report that the cosmic-noise attenuation at College, Alaska, persisted for about 60 hours after the arrival of the particles produced by the flare on July 16. Even for events of much less intensity than this one the enhanced absorption is often present for times considerably longer than this.

At the time cosmic-noise attenuation disappeared at College there remained very great fluxes of particles above 85 Mev energy at Resolute Bay. The ionization produced by these flare protons was 10 times greater than that due to the normal cosmic rays at  $6 \text{ g cm}^{-2}$  depth. Furthermore, it is likely that the  $E^{-4}$  integral spectrum may still be applied at this time even up to energies of 300 or 400 Mev. It is, therefore, not possible to hold the view that the disappearance of the cosmic-noise absorption effects at College is due to exhaustion from the storage region of higher-energy particles with only particles of lower energy present at higher geomagnetic latitudes. Apparently there is some mechanism that drastically reduces the particle flux at College while allowing it to remain farther to the north.

During the August 22, 1958, event at a time when the ionization at  $6 \text{ g cm}^{-2}$  depth due to particles originally of energy above 85 Mev was 1.5 times that due to normal cosmic radiation, there was 8 db of cosmic-noise attenuation at Fort Churchill. It was shown that most of the

ionization producing the attenuation was due to protons having energy from 20 to 80 Mev in an  $E^{-4}$  integral spectrum. This means that for the present event very effective removal of particles over this low-energy region must be taking place above College. No balloon measurements at that place are available to decide what the effect on particles above 80 Mev was at this time (60 hours after the first arrival of the flare particles).

We have no interpretation for these particular observations, but it may be pertinent to note that the magnetic storm followed the July 16 solar flare by only 19 hours. The removal of the low-energy particles could then be due to an enhancement of the geomagnetic (or interplanetary) field, hence increased rigidity cutoffs at the earth's surface due to some effect of the solar gas cloud.

8. The times at which the large influx of particles occurs at geomagnetically forbidden latitudes coincide very closely with an anomalous effect which has been observed in neutron monitors at latitudes south of Resolute [Wilson, Rose, and Pomerantz, 1960]. This effect is a sudden increase of 5 to 10 per cent that follows the steep-descent portion of the Forbush decrease and persists for 3 to 4 hours. Since the flare particles are present in great intensity above Resolute without their effects being seen in the neutron monitor there, it is doubtful that this increase is due to these low-energy cosmic rays. We suppose that it must be due to a temporary return of some of the high-energy cosmic rays earlier excluded by the Forbush decrease and that this effect is associated with the magnetic-storm feature that lets in the low-energy flare particles at forbidden latitudes. This storm feature is clearly complex, since it does not readmit high-energy particles at Resolute Bay.

9. Data now available on the low-energy solar cosmic-ray emission permit estimating the sun's production rate of protons in the energy region 30 to 300 Mev. Because of the frequency of this type of event, and since it is currently believed that the emission of particles from the region of the sun continues for several days, this rate is several orders of magnitude higher than the production rate for very high-energy cosmic rays such as occurred on February 23, 1956.



aged over the current sunspot maximum (1956-1960) the flux arriving over the polar regions is  $10^8$  protons/(cm<sup>2</sup>-year) having energy above 100 Mev. Then, assuming the par- emission to take place over an entire solar sphere, the corresponding yearly production by the sun is of the order of  $10^{35}$  protons. For a complete sunspot cycle the rate would probably be about half this figure, although the accuracy of the maximum rate just given is probably not as good as a factor of 2.

There is direct evidence that, during the event which began on August 22, 1958, the  $E^{-2}$  differential energy spectrum derived at that occasion is valid down to 20 or 30 Mev (Anderson, Arnoldy, Hoffman, Peterson, and Pickler, 1959; Rothwell and McIlwain, 1959). Furthermore, riometer data make it quite certain that during many of these events large numbers of particles of energy this low and perhaps lower must be present (Reid and Collins, 1959; Hultquist, 1959). Applying the  $E^{-2}$  inter-law to all the low-energy cosmic-ray events, the yearly flux over 30 Mev arriving over the polar caps is  $10^{10}$  protons/(cm<sup>2</sup>-year) and the corresponding solar production rate is  $10^{37}$  protons/year. The total kinetic energy carried by the particles is about  $5 \times 10^{30}$  ergs/year, and if the rest-mass energy is included, the rate of energy loss by this process is  $2 \times 10^{34}$  ergs/year. The rest-mass energy is probably quickly replenished by electrostatic pickup of hydrogen from interplanetary gas.

Very little detailed information is available on the proton energy spectrum below about 30 Mev. During the time observations were being made on the flare radiation no substantial changes in the average particle energy were noted. In particular the average energy was not found to increase, and it is more likely that it decreased. It is therefore concluded that, after leaving the region of the sun, the particles in the inner solar system experience no further acceleration that produces substantial changes of energy on the time scale of days.

Furthermore, the flux of flare particles has nearly disappeared at the earth 11 days after the flare, and on the twelfth day observations by Neher (private communication) established that the particles were no longer present. The

lifetime of the particles in the inner solar system is, then, of this order of time or less, so that any accelerating mechanism operating on a longer time scale is ineffective in producing energy changes. Therefore, the low-energy protons from the sun do not directly contribute in any substantial way to the high-energy cosmic-ray flux at the earth. Conversely, this means that the high-energy cosmic-ray spectrum [McDonald, 1959] is nonlocal to a high degree in the sense that the inner solar system (except for the sun itself) is not able to confine or accelerate protons effectively over the energy region 20 Mev to greater than 20 beV (when results from the February 23, 1956, flare are also used).

With present knowledge of the composition and energy spectrum of the flare particles, some rather limited statements can be made about the possibility that these solar emissions are suitable injection material for proposed galactic acceleration processes.

Considering first a mechanism that gives a particle energy gain per collision proportional to  $B$  in the first order [Davis, 1956], where  $B$  is the velocity of the 'wall' divided by the velocity of light, the injection energy would be perhaps as low as 10 Mev, although the exact value would depend critically on the detailed properties of the interstellar region occupied by the particles. Thus in the case of the sun perhaps all the copious output of nucleonic particles from solar flares would be accelerated. Then if the sun is somehow typical of the other stars in the galaxy the injector strength thus available could be enough numerically to maintain the high-energy cosmic-ray beam. As a very rough illustration of this point, if the estimated yearly rate of production of solar protons above 10-Mev energy, taken as  $10^{30}$ , is applied to all  $10^{11}$  stars in the galaxy, a yearly galactic rate of  $10^{40}$  protons is arrived at. Then, taking the galactic volume as  $2 \times 10^{67}$  cm<sup>3</sup> and the cosmic-ray density as  $10^{-10}$  cm<sup>-3</sup>, the number of cosmic rays turns out to be  $2 \times 10^{67}$ . Thus the injection rate on this basis permits an average lifetime of at least  $10^7$  years for the high-energy cosmic radiation.

The lack of helium at the higher rigidities where measurements presently apply throws serious doubt on the possibility that these abundant solar particles bear very directly on the

question of high-energy galactic cosmic-ray origin. It will no doubt be necessary to have rocket or satellite measurements of the very low  $\alpha$ -particle rigidity region in order to make a conclusive statement on this point, although the riometer data seem to be adequately explained by the flare radiation having only a proton component.

Considering the original Fermi acceleration mechanism, it seems clear that these solar particles are not a suitable injection material, since the injection energy is very high for helium and measurements clearly show the absence of fast helium from solar flares.

11. The production of low-energy cosmic radiation by the sun is predominantly a large flare effect. This fact is best seen from the riometer data [Leinbach and Reid, 1959], where over 80 per cent of these solar events are associated with flares of importance 3 or 3+. The far more frequently occurring class 1 and class 2 flares have produced marked cosmic-noise attenuation effects only on a very few occasions. There is the possibility that small but significant fluxes of protons are accelerated by small flares and remain undetected by this radio technique. For the rather characteristic  $E^{-4}$  integral spectrum found in these events, particle detectors carried by balloons to atmospheric depths of 10 g cm<sup>-2</sup> or less offer a means about 10 times as sensitive for detecting the presence of these protons. Therefore we have examined the data from our 44 high-altitude flights made in or north of the auroral zone during 1957, 1958, and 1959 for small, sporadic counting-rate effects occurring at times other than known solar cosmic-ray events. The ion-chamber records were analyzed for pulse-rate increases larger than about 3 per cent averaged over 10 ion-chamber pulses, ordinarily about 10 minutes. No such effects could be found after account was taken of changes in the balloon's altitude, daily variation of the galactic cosmic rays, and auroral zone X ray fluxes. This analysis was made over more than 700 hours of high-altitude data during which time there appeared on the sun about 220 flares of importance 1 or 1+ and 22 flares of importance 2 or 2+. There were also two class 3 flares which produced no effects in the particle detector.

During the course of this high-latitude bal-

loon program we have obtained information on four separate low-energy solar cosmic-ray events, all of which were associated with importance 3 or 3+ flares. In addition to the August 22, 1958, occurrence, measurements of the one in progress during September 1, 1959, have been described previously [Anderson, 1960]. Some brief observations were obtained on particles that appeared after the large flare of August 26, 1958, but their chief usefulness has been to verify the more significant Explorer IV results [Rothwell and McIlwain, 1959] of that occasion. The cosmic-ray aspects of the great July 16, 1959, flare have been described in the present paper. A total of 150 hours of high-altitude observations has been obtained of this phenomenon, mainly contributed by the last event. In addition to these there is the important negative result that on August 2, 1957, during a marked polar-cap riometer event, no proton effects above 100 Mev were present.

The fact that only very large flares produce energetic nucleonic particles suggests that very great amounts of energy must be available to create the special conditions necessary for acceleration, and perhaps the emission as well, of protons in the 10- to 300-Mev energy range. Severny [1959] has reported that the sunspot magnetic fields in the region where the large flare of August 22, 1958, occurred were of simpler configuration and reduced in strength after this flare, thus implying a large energy release. However, Howard, Cragg, and Babcock [1959] at Mount Wilson Observatory, observed no magnetic field rearrangement or destruction in the sunspot region where the great flare of July 16, 1959, appeared, making it doubtful that the flare energy was derived so close to the photosphere.

*Acknowledgments.* Our use of the Resolute Bay site in Canada was made possible by the efforts of Dr. D. C. Rose of the National Research Council and Mr. G. W. Rowley of the Department of Northern Affairs. The Royal Canadian Air Force provided the necessary airlift of the equipment and personnel. The U. S. Weather Bureau assisted in transporting the helium supply. The personnel at the Resolute Bay Ionospheric Station operated by the Canadian Department of Transport, and Mr. W. Fedorak, Officer in Charge, provided us with much field support and congenial companionship. The U. S. Office of Naval Research assisted the project in many administrative ways, particularly

gh Mr. G. W. Bowers, the Minneapolis representative.

message from the High Altitude Observatory Boulder, Colorado, warning us of the possibility of solar proton effects on the basis of the large and radio noise storm that occurred on July 14 as of much assistance in planning our operation. Access to the Fort Churchill riometer operated by the Canadian Defence Research Board permitted us to follow the course of the two major solar proton events in July while waiting for the airlift to Resolute Bay.

J. D. Johnston of the Canadian National Research Council provided us with invaluable help in preparing the balloon launchings and in monitoring the telemetered signals. Finally, we thank the third member of the University of Iowa party, Robert D. Lamb, for his many hours of help, particularly his personal efforts in the critical job of anchoring the balloon just before release.

The work described here was supported by the National Science Foundation and was part of the United States' participation in International Geophysical Year 1959.

#### REFERENCES

- Erson, K. A., Soft radiation events at high latitude during the magnetic storm of 29-30 August 1957, *Phys. Rev.*, **111**, 1397-1405, 1958a.
- Erson, K. A., Ionizing radiation associated with solar radio noise storm, *Phys. Rev. Letters*, **3**, 335-337, 1958b.
- Erson, K. A., Balloon observations of X rays from the auroral zone I, *J. Geophys. Research*, **65**, 549-564, 1960.
- Erson, K. A., R. Arnoldy, R. Hoffman, L. Peterson, and J. R. Winckler, Observations of low-energy solar cosmic rays from the flare of 22 August 1958, *J. Geophys. Research*, **64**, 1133-1137, 1959.
- Erson, R. R., and R. G. D'Arcy, Observations of solar flare radiation at high latitude during the flare of July 10-17, 1959, *Phys. Rev. Letters*, **3**, 392, 1960.
- Charakhch'yan, A. N., V. F. Tulinov, and T. N. Charakhch'yan, *Cosmic Rays Emitted by the Sun*, preprint, Lebedev Physical Institute, January 1960.
- Davis, L., Modified Fermi mechanism for the acceleration of cosmic rays, *Phys. Rev.*, **101**, 351-358, 1956.
- Enemark, D., Balloon-borne circuits sort high altitude cosmic rays, *Electronics*, **32** (35), 52, 1959.
- Erickson, W. C., Radio emission in the outer corona, *Phys. Rev. Letters*, **3**, 365-367, 1959.
- Fermi, E., On the origin of the cosmic radiation, *Phys. Rev.*, **75**, 1169, 1949.
- Freier, P. S., Energy spectrum of protons from the sun, Midwest Cosmic Ray Conference, Iowa City, Iowa, October 30-31, 1959.
- Freier, P. S., E. P. Ney, and J. R. Winckler, Balloon observations of solar cosmic rays on March 26, 1958, *J. Geophys. Research*, **64**, 685-688, 1959.
- Howard, R., T. Cragg, and H. W. Babcock, Magnetic field associated with a great solar flare, *Nature*, **184**, 351, 1959.
- Hultquist, B., On the interpretation of ionization in the lower ionosphere occurring on both day and night side of the earth within a few hours after some solar flares, *Tellus*, **11**, 332-343, 1959.
- McDonald, F. B., Primary cosmic ray intensity near solar maximum, *Phys. Rev.*, **116**, 462, 1959.
- Ney, E. P., J. R. Winckler, and P. S. Freier, Protons from the sun on May 12, 1959, *Phys. Rev. Letters*, **3**, 183-185, 1959.
- Reid, G. C., and C. Collins, Observations of abnormal VHF radio wave absorption at medium and high latitudes, *J. Atmospheric and Terrest. Phys.*, **14**, 63, 1959.
- Reid, G. C., and H. Leinbach, Low-energy cosmic-ray events associated with solar flares, *J. Geophys. Research*, **64**, 1801-1805, 1959.
- Rothwell, P., and C. E. McIlwain, Satellite observations of solar cosmic rays, *Nature*, **184**, 138, 1959.
- Severny, A. B., *Publ. Crimean Astrophys. Observatory*, **22**, 1959.
- Wilson, B. G., D. C. Rose, and M. A. Pomerantz, The unusual cosmic ray events of July 17-18, 1959, *Can. Natl. Research Rept.*, January 1960.

(Manuscript received May 31, 1960.)





## Comparison of Solar Cosmic Rays Injection Including July 17, 1959, and May 4, 1960

K. G. McCracken and R. A. R. Palmeira<sup>1</sup>

*Physics Department and Laboratory for Nuclear Science  
Massachusetts Institute of Technology  
Cambridge, Massachusetts*

**Abstract.** Using neutron monitor data obtained at high latitudes, it is shown in Part 1 that cosmic-ray particles of rigidity  $> 1$  bv were produced by the solar flare that occurred at 2115 UT on July 16, 1959. The exponent of the integral rigidity spectrum of the radiation was  $\leq 0$ . The temporal dependence of the flare radiation was unlike that of earlier flare effects, the intensity requiring  $\sim 7$  hours to reach its maximum value. No impact zones were observed. In Part 2, meson data obtained at the time of the flare effect starting at 1031 UT on May 4, 1960, are presented. In Part 3, by considering the above-mentioned flare effects, and those observed previously, it is shown that the time scale of the cosmic-ray flare effect is a highly variable quantity, varying by a factor as great as 36 from event to event. The flares responsible for the four flare effects with the shortest time scales occurred on the western solar limb; those resulting in the flare effects with the longest time scales occurred near the center of the solar disk. In Part 4 the above facts are shown to support the theoretical model proposed by Pidginton.

### THE FLARE EFFECT OF JULY 17, 1959

**Production.** Over a period of 23 years, six events have been observed in which an intense flare was followed within a few hours by a rapid enhancement in the cosmic-ray intensity at the surface of the earth [Dorman, 1957; McDaniel, 1959]. These six events showed a number of marked similarities, and it seemed reasonable to list the characteristics of solar flares as follows:

The cosmic-ray intensity increases rapidly to a maximum intensity occurring within about 100 minutes of the observation of the maximum magnetic radiation from the flare. Once maximum is reached, the intensity commences a smooth decline to the pre-event level. The time required to attain maximum intensity is considerably less than that required for the intensity to return to the pre-event level.

The integral rigidity spectrum of the produced cosmic radiation is very steep, falling off approximately as  $(\text{rigidity})^{-7}$ .

The event occurs at roughly the same time at all points on the earth. The small differences in onset time ( $\sim 10$  minutes) that do

occur can be explained as follows [Lüst and Simpson, 1957]. The higher-rigidity particles arrive at the earth first, and appear to come from within a relatively small solid angle in the direction of the sun. Thus, at first, only stations within several limited regions (impact zones) on the earth's surface observe a cosmic-ray enhancement. With the passage of time, lower-rigidity particles arrive, the radiation becomes isotropic, and all stations for which the rigidity cutoff is low enough observe an intensity enhancement.

In what follows, we consider an intensity enhancement which followed soon after an intense 3+ flare, but which, while fulfilling criteria (b) and (c), failed to fulfill criterion (a). In particular, the period of increasing intensity was of at least 6 hours' duration, maximum intensity occurring some 8 hours after the maximum output of visible light from the flare. We interpret this event as being a genuine case of the production of relativistic particles at the sun, and suggest that the atypical time behavior was the result of a disordered magnetic field between the earth and the sun, the cosmic-ray particles being required to diffuse through this field.

**Observations.** The cosmic-ray neutron data obtained at Mawson, Antarctica, and Herst-

<sup>1</sup> Leave from Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil.

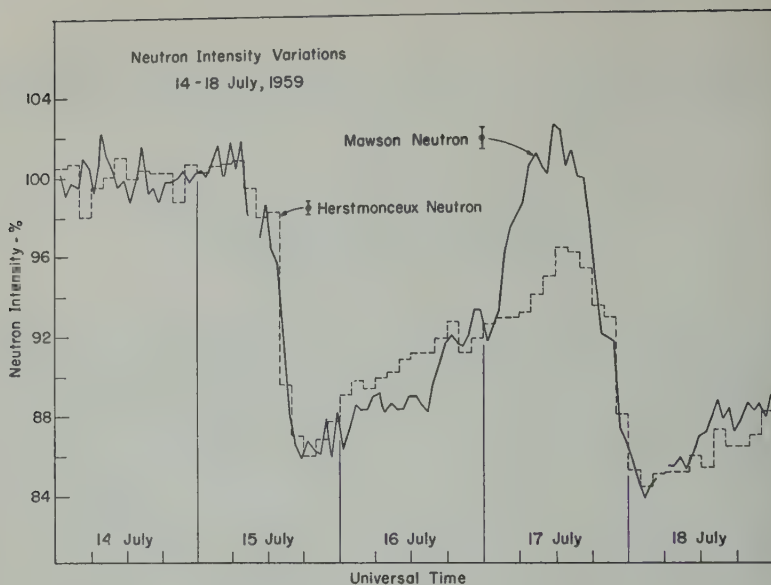


Fig. 1. Neutron monitor data from Mawson and Herstmonceux for the period July 14-18, 1959.

monceux, England, during the period July 14-18, 1959, are displayed in Figure 1. The Herstmonceux data show that two very large Forbush-type decreases occurred, one early on July 15, the other late on July 17. It can be seen that the Mawson curve tracks with the Herstmonceux curve quite well except during the period 02 to 16 UT on July 17. This failure to track can be seen to be due to a sudden change in slope in the Mawson curve, the intensity at that station increasing by about 8 per cent within a period of about 4 hours.

The neutron data obtained at 31 observatories were examined, and the data from many of them were plotted as in Figure 1. Separate graphs were prepared for the longitude zones centered on Europe, the Americas, and Australia, this procedure being adopted as it is known that during the recovery phase of a Forbush decrease there are differences in the behavior of the cosmic-ray intensity at different longitudes. It was seen from these graphs that the intensity changes at a number of observatories were similar to those at Herstmonceux: that is, a typical recovery from the Forbush decrease on July 15 preceded the Forbush decrease on July 17.

The data from another group of stations failed to track with those from the first group. This failure to track was invariably due to a sudden

intensity increase at about 02 UT on July 17, the curves being very similar to those for Mawson. The increment was not present in the data from these observatories. This fact by itself has already prompted a number of investigators to speculate that the neutron enhancement was due to solar-produced radiation [Carmichael and Steljes, 1959; Bailey and Pomerantz, 1960]; however, the fact that the time changes during the event were unlike those of earlier flare effects has thrown some doubt on this interpretation. It is clear that we must investigate other characteristics of the anomalous increase to see whether they support the hypothesis of solar production.

In Table 1, some of the stations whose data we have examined are listed in order of increasing cutoff rigidity  $N$ . The cutoff rigidities calculated by Quenby and Webber [1959] were employed. It is immediately clear that every station for which  $N$  was 1.13 bv or less observed an enhancement, but it was not seen by any station for which  $N$  was 1.15 bv or more. This suggests that the rigidity spectrum of the enhanced radiation was very steep. Furthermore, it can be seen from the table that, if the enhancement was observed, it commenced between 01 and 02 UT irrespective of the longitude of the station.

To obtain some quantitative data on the



TABLE 1. Details of the Anomalous Cosmic-Ray Events Observed at High-Latitude Stations during July 17, 1959

| Station                  | Cutoff Rigidity, bv | Amplitude, % | Onset Time, GMT |
|--------------------------|---------------------|--------------|-----------------|
| Anomalous event seen     |                     |              |                 |
| Uppsala                  | <0.10               | 5.0          | 01-02           |
| Herstmonceux             | <0.10               | 5.0          | About 04        |
| Bergen                   | 0.11                | 5.0          | 02-04           |
|                          | 0.75                | 5.0          | 03-04           |
|                          | 0.94                | 4.7          | About 04        |
|                          | 0.96                | 4.0          | 03-05           |
| Deep River               | 0.97                | 4.5          | 01-02           |
| Sulphur Mt.              | 0.98                | 11.0         | 01-02           |
| Washington               | 1.03                | 4.5          | 02-04           |
| Chicago                  | 1.13                | 4.5          | 02-04           |
| Anomalous event not seen |                     |              |                 |
| Chicago                  | 1.15                |              |                 |
| Chicago                  | 1.54                |              |                 |
| Chicago                  | 1.70                |              |                 |
| Chicago                  | 1.77                |              |                 |
| Herstmonceux             | 2.25                |              |                 |
| Bergen                   | 2.42                |              |                 |

Anderson, Rose, and Pomerantz [1960].

Anderson, J. G., O. R. Santochi, J. C. Anderson, J. R. Manzano, and J. R. Manzano, to be published.

of the rigidity spectrum the data from Herstmonceux, Uppsala, and Bergen have been compared. The fact that these stations are quite close to one another gives us confidence that, if there are errors in the assigned cutoff rigidities, they will be of approximately the same magnitude, and that therefore the ratio of the rigidities appropriate to any two stations is approximately correct. It is also known that during Forbush decreases, the percentage variation in intensity at these three stations is approximately equal, and consequently, as Herstmonceux did not see the enhanced radiation, it was possible to remove the variations due to the Forbush decrease from the Bergen and Uppsala data by subtraction of the percentage variations observed at Herstmonceux from the variations observed at the other two stations. The results of these calculations are shown in Figure 2. Averaging over the period during which Bergen saw the enhanced radiation, the amplitude to be 4.5 per cent at Bergen and 0.3 per cent at Uppsala. Assuming

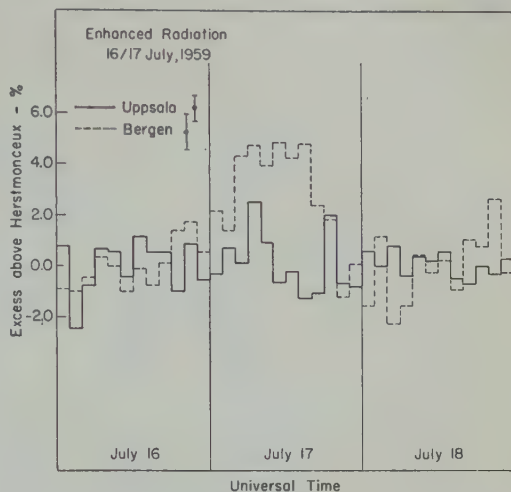


Fig. 2. Enhanced radiation at Bergen and Uppsala measured as an excess above the cosmic-ray background as measured at Herstmonceux.

a power law, with exponent  $-\beta$ , for the integral rigidity spectrum of the enhanced radiation, the above values set a lower limit of 8.1 for  $\beta$ . An upper limit cannot be set on account of the statistically insignificant increase at Uppsala. The fact that the atmospheric cutoff is in the vicinity of 1 bv means that the effective cutoff at Bergen may be greater than the value used (0.94 bv), and consequently  $\beta = 8.1$  may be an underestimate of the lower limit for  $\beta$ .

A further check on the value of the exponent was obtained by comparison of the enhancements observed at the mountain stations of Mt. Washington (1910 meters altitude, 1.03 bv vertical cutoff rigidity) and Sulphur Mountain (2280 m, 0.98 bv). It was known from the stations in the Americas that did not observe the increment (Chicago, Berkeley, Buenos Aires) that at these longitudes the recovery of the Forbush decrease during July 17 approximated quite well to a linear extrapolation of the recovery during July 16. Using this approximation, the intensity of the enhanced radiation at stations in the Americas was derived, the data from three stations being given in Figure 3. The data from Deep River and Sulphur Mountain were used to derive the attenuation length of the enhanced radiation, a value of  $89 \text{ g cm}^{-2}$  being obtained. This in itself suggests that the enhancement was not merely a modulation of the galactic radia-

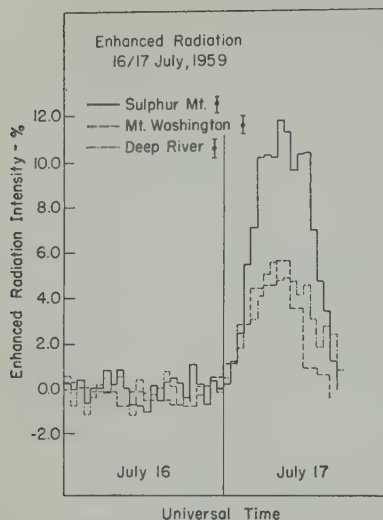


Fig. 3. Enhanced radiation at Sulphur Mt., Mt. Washington, and Deep River measured as an excess above the cosmic-ray background estimated for each station.

tion, as the value of the attenuation length applicable to the galactic radiation is  $138 \text{ g cm}^{-2}$ . Taking the value  $89 \text{ g cm}^{-2}$  for the attenuation length, the Sulphur Mountain observations were used to estimate the enhancement that would have been observed at an altitude of 1910 m at Sulphur Mountain. This value, and the data obtained at the same altitude at Mt. Washington, gave a spectral exponent of  $9.3 \pm 2.0$ .

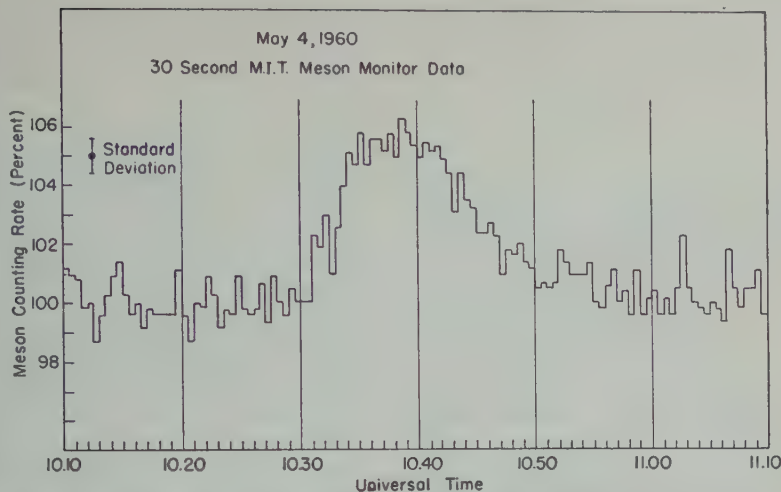
Thus the criteria (b) (steep rigidity spectrum) and (c) (approximate simultaneity) listed in the Introduction were fulfilled by the intensity enhancement observed on July 17, 1959. As the enhancement, and the associated ionospheric effects [Bailey and Pomerantz, 1960; Reid and Leinbach, 1959] indicative of the arrival of lower-rigidity particles occurred soon after the observation of a 3+ flare (at 2115 UT on July 16 at a point  $30^\circ$  west of the central solar meridian [National Bureau of Standards, 1959], there is little doubt that the sea-level enhancement was of the nature of a solar flare increase.

*Isotropy of the flare radiation.* During previous examples of the solar flare effect, the cosmic-ray intensity within the '9 o'clock impact zone' was greater than that in the 'background zone' by a factor of 2 or more [Lüst and Simpson, 1957; Firor, 1954]. This inequality persisted

for an hour or so after the maxima of the effects, and was consistent with a short-lived source in the direction of the sun. Assuming an apparent source of  $40^\circ$  in latitude and  $80^\circ$  longitude, centered on the sun, it was found that both Mawson and Bergen would have been within the 9 o'clock zone during the period from 06 UT on July 17. The event amplitude at these two stations and the detailed shapes of the increasing intensity phases were not significantly different from those observed at sea level at other points on the earth's surface, leading to the conclusion that impact zones were not formed. It therefore seems that by 03 UT, about 2 hours before maximum intensity was attained, the radiation was already isotropic. The data are not sufficiently accurate to determine whether the radiation was isotropic before 03 UT. The arrival of flare radiation at Thule, Resolute, and Churchill also requires isotropy by about 03 UT.

*Temporal dependence of the enhanced radiation.* The maximum emission of H $\alpha$  radiation occurred at 2230 UT on July 16, and the maximum ionospheric effect indicative of the arrival of corpuscular radiation at the earth commenced at 2250 UT. The Sulphur Mountain and Deep River data, that is, the neutron data with the best statistics, show the first noticeable flare effect between 01 and 02 UT, or  $\sim 2$  hours after the commencement of the ionospheric effect. The neutron data from Sulphur Mountain and Deep River indicate that the flux of solar cosmic radiation continued to increase until about 05 UT. To understand the difference in the times of onset of the ionospheric and neutron effects, we must consider the sensitivity of the two types of measurement to low-rigidity solar protons.

Using the primary proton spectrum found by McDonald and Webber [1959], and the dependence of neutron intensity upon latitude observed by Rose, Fenton, Katzman, and Simpson [1959], the specific yield function [Fonger, 1953] for 1-bv protons was calculated, and used to estimate that it would require a flux of  $\sim 15$  protons  $\text{ster}^{-1} \text{sec}^{-1}$  ( $100 \text{ Mv}^{-1}$  in the rigidity range 1 to 1.1 bv to produce a 0.5 per cent, and barely detectable, increase in the neutron counting rate at a station at a high latitude. On the other hand, Reid and Leinbach [1959] estimate that a riometer can readily detect a flux of



The solar flare cosmic-ray increase on May 4, 1960, as observed by the MIT meson detector.

$\text{m}^{-2} \text{sec}^{-1}$  if the protons are in the rigidity range 0.05 to 0.3 bv. Using this value, and the integrated rigidity spectrum of the solar radiation (going as the reciprocal of the rigidity power of the rigidity above 0.3 bv), we estimated that this threshold flux for the flare corresponds to  $1 \text{ proton m}^{-2} \text{sec}^{-1} \text{ster}^{-1} (\text{Mv})^{-1}$  at 1 bv, that is, a flux which is a factor of 15 below the neutron monitor threshold. These considerations show that there is no inconsistency between the riometer and neutron monitor. A plausible reason for the difference in time of onset being that a riometer can detect much smaller flux of solar protons than a neutron monitor. It is concluded that the solar radiation first arrived at the earth at 2250 UT, the flux continuing to increase until a maximum was reached at about 05 UT on July 17, 1959, 8 hours after the maximum emission of the flare. This slow rise in cosmic-ray intensity is in marked contrast to the abruptness of the rise during previous flare effects.

#### THE FLARE EFFECT OF MAY 4, 1960

While this paper was in the final stages of preparation, another flare effect was observed by ground-level cosmic-ray detectors. As this effect adds considerably to the experimental material to be considered in the next section, we thought it advisable to rewrite a part of this section in order to include the new observations. The discussion in the next section will be

largely concerned with the time scales of flare effects, we will use the meson data obtained at MIT for the sole purpose of establishing the time scale of the May 4 event.

Figure 4 displays the sum of the counting rates observed by the three large meson telescopes at MIT. The instruments have been described elsewhere [Palmeira and Williams, 1958], and it suffices to point out that there is now 7.5 cm of lead between the upper and lower scintillators. Each telescope has a counting rate of about  $300 \text{ sec}^{-1}$ .

The cosmic-ray flare effect can be seen to have commenced at about 1031 UT. This would place MIT in the background zone [Firor, 1954].<sup>2</sup> A preliminary report [National Bureau of Standards, 1960] indicates that a solar flare of im-

<sup>2</sup>Note added in proof. The result that MIT lay in the background zone during the May 4, 1960 event was based on the assumption of an apparent source in the direction of the sun. However, analysis of the observations made at many points on the earth's surface has led to the conclusion that MIT did in fact lie in the 9 o'clock (primary) impact zone, and that the apparent source of solar cosmic radiation was displaced from the direction to the sun by a considerable angle (approximately  $70^\circ$  to the west). This result does not alter the conclusions of this paper for, allowing a factor of two for the ratio of the time scales applicable to the background, and primary impact zones, it is still true that the May 4, 1960 event had the shortest time scale of any of the observed flare effects.



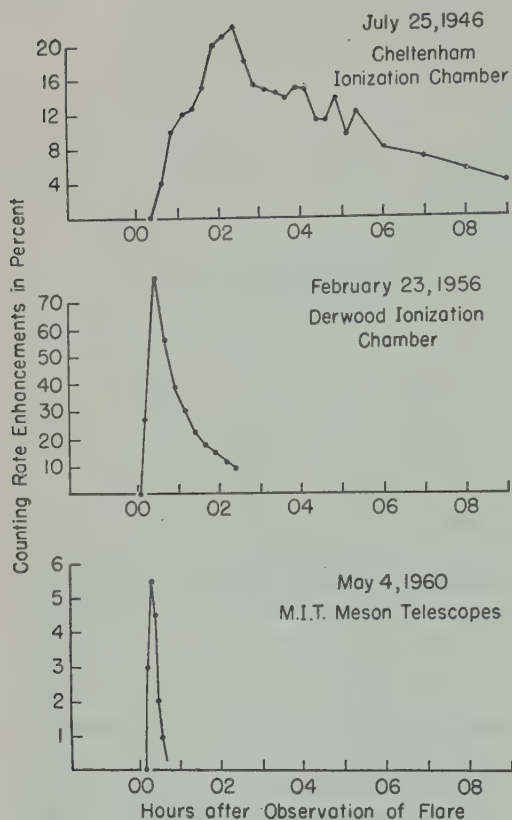


Fig. 5. Meson data for the solar flare cosmic-ray increases of July 1946, February 1956, and May 1960.

portance 2 was observed to start at about 1020 UT in the vicinity of a sunspot group 90° west of the central solar meridian. It was not followed by a Forbush decrease. At the time of occurrence of the flare, the terrestrial cosmic-ray intensity was recovering from a Forbush decrease which commenced on April 30.

### 3. THE TIME SCALES OF FLARE EFFECTS

In Figures 5 and 6 are displayed the meson and neutron data obtained at background zone stations during the July 1946, February 1956, July 1959, and May 1960 flare effects [Forbush, Stinchcomb, and Schein, 1950; Forbush and Burke, 1956; Rose and Katzman, 1956]. Visual comparison of these graphs shows that the time scales of the events differ markedly; this will now be established in a more quantitative manner. It is known that the observations of any

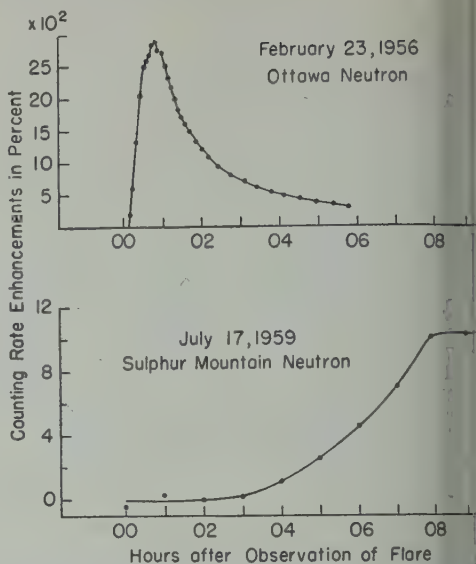


Fig. 6. Neutron data for the solar flare cosmic-ray increases of February 1956 and July 1959.

single flare effect can have time scales differ by a factor of approximately 2, depending on the type of instrument used, the cutoff rigidities of the stations, and the type of flare effect in which the stations lay (impact, background, or forbidden polar zone). Consequently, in order to avoid uncertainties in the comparisons between different events, we have, wherever possible, compared data obtained by similar instruments, situated at stations in the background zones where the cutoff rigidities are roughly the same.

For convenience, the interval between commencement of the intensity enhancement and the time at which the maximum intensity was attained will be called the rise time of the event. The interval over which the enhancement subsequently decreased by half will be called the 50 per cent decay time.

Comparing the MIT, Derwood, and Cheltenham meson data (Fig. 5), the rise times of 1960, 1956, and 1946 events are found to be minutes, <30 minutes, and ~120 minutes, the 50 per cent decay times to be ~8 minutes, ~40 minutes, and ~180 minutes, respectively. That is, the time scales of the rising intensity phases of these three events are in the ratio 1:4:17; the time scales of the decreasing intensity phases are in the ratio 1:5:22. C

the Ottawa and Sulphur Mountain neutron (Fig. 6), the rise time of the 1956 and 1959 events are found to be  $\sim 45$  minutes and 1 hour; that is, the time scales are in the ratio 1:2. The fact that a Forbush decrease occurred on July 17, 1959, prevents a determination of the 50 per cent decay time of the event.

Combining the above data, it is found that the times of the May 1960, February 1956, July 1946, and July 1959 events are in the ratio 1:3:6. Furthermore, the evidence suggests that the 50 per cent decay times exhibit approximately the same ratio. That is, whereas the time of flare effects can differ from event to event, the rise time, measured relative to the 50 per cent decay time, is a constant. This fact seems to be an important clue to the nature of the mechanism responsible for the observed temporal dependence of flare effects.

For completeness, the time scales of the February 1942, March 1942, and November 1949 events [Dorman, 1957] have been compared with those of the May 1960 event. The values for the rise times are listed in Table 2. For the February 1942 event, the rise time had to be determined from data obtained at a station in the impact zone (Cheltenham). To permit comparison with the other determinations, the value so found was multiplied by a factor of 1.5. The data obtained during the February 1956 event suggesting that this is a reasonable value for the time scale of an event observed in the background zone when measured relative to the time scale of the same event as observed in the main impact zone). It is found that in all cases the ratio of the rise time to the 50 per cent decay time is approximately constant.

#### 4. DISCUSSION

If a solar flare effect, the Forbush-type decrease, and the 11-year cycle of cosmic-ray intensity are responsible for introducing the great changes into the terrestrial cosmic-ray intensity. An understanding of these three effects will take us a long way toward an understanding of the whole field of cosmic-ray variations. Clearly any satisfactory theory must explain the observations obtained when two or more of the effects are superposed. The super-

position of a flare effect upon one or both of the other two is especially valuable, as the time of flight of the cosmic radiation and the degree of isotropy upon arrival at the earth are indicative of the electromagnetic conditions in interplanetary space, and therefore permit a test to be made of any models that may be proposed to explain the Forbush-type decrease and the 11-year cycle of intensity change. Such a test will now be applied to a model suggested by theoretical studies.

The model we will examine is essentially that proposed by Piddington [1958] and, in part, by others [Morrison, 1956; Parker, 1958; Cocconi, Greisen, Morrison, Gold, and Hayakawa, 1958]. It is suggested that during a period of solar activity the sun is surrounded by an extended region containing a disordered magnetic field ( $10^{-5}$  gauss), generated in a manner to be described later. When a large solar flare occurs in an active sunspot region, matter is ejected, which, on account of its high electrical conductivity, carries part of the sunspot magnetic field with it. Within the cloud of matter, there will be chaotic material motions, and the magnetic field will be fairly strong ( $\gtrsim 10^{-4}$  gauss) and disordered. However, there will be lines of force connecting the cloud to the sunspot magnetic field, and these lines will be relatively ordered, and radial, as shown in Figure 9. The radial magnetic field remains rooted in the sunspot group, and, consequently, the rotation of the sun causes it to become twisted as shown in Figure 7 some days after the occurrence of the flare. In an active sunspot group, there is a more or less continual outflow of matter, which, though not as energetic as the matter ejected in a large flare, nevertheless establishes an ordered field stretching out from the sun. With the passage of time, instabilities destroy the ordered nature of the radial field, the disordered remnants contributing to the general disordered field postulated previously.

It has been shown that mechanisms based upon particle diffusion through disordered magnetic fields might be responsible for the 11-year and Forbush-type variations [Morrison, 1956; Parker, 1958]. In the present model, it is proposed that the weak but extensive general disordered field is responsible for the 11-year variation, while the stronger, chaotic fields in a cloud

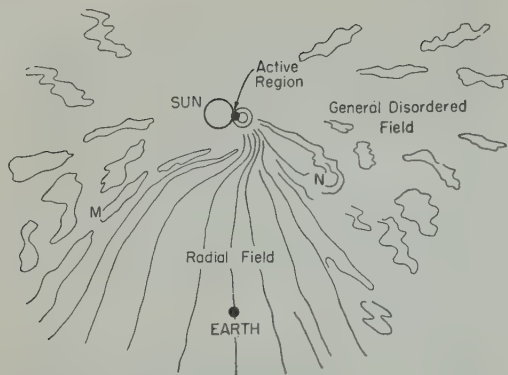


Fig. 7. Case A of the proposed model for the magnetic fields in the sun-earth region. In region *N*, matter ejected by the sun is stretching out the solar magnetic fields, thereby forming an essentially radial field. In region *M*, instabilities are destroying the radial field.

of matter ejected by a large solar flare are responsible for the Forbush-type decrease.

We shall now consider the consequences predicted by this model in the event of sudden production of cosmic radiation in a solar flare. Consider the three cases that follow:

*Case A* (Fig. 7): An active group has passed across the solar disk and is now on the western limb. Matter ejected by the group has established an essentially radial field enveloping the earth. If a solar flare did occur while the active center was near central meridian passage (c.m.p.), the highly chaotic matter and fields have by now passed the earth, and the field near the earth is essentially radial.

In such a case, cosmic rays produced by a solar flare in the active region would gain almost immediate access to the earth by spiraling along the lines of force of the radial field. The small irregularities that would undoubtedly be present in the radial field would alter the pitch angles of the cosmic radiation, thereby producing small time delays and dispersion of the high- and low-energy parts of the cosmic-ray spectrum (as the change in pitch angle would be an energy-dependent function). The first particles to arrive at the earth would be those moving roughly parallel to the radial magnetic field, and, consequently, there would be an apparent source in the direction of the sun during the early stages of the cosmic-ray enhancement. At later times,

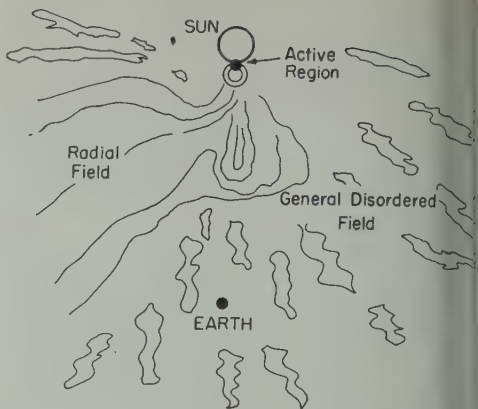


Fig. 8. Case B of the proposed model for magnetic fields in the sun-earth region.

radiation scattered by the irregularities of the radial field would cause the radiation at earth to become roughly isotropic.

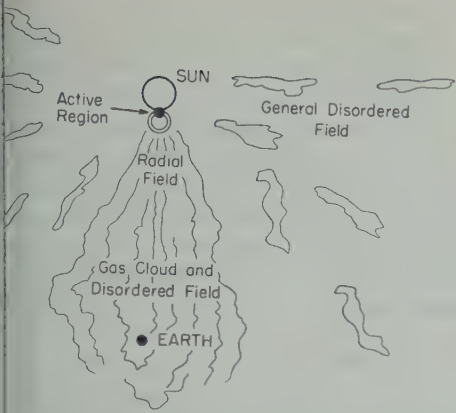
*Case B* (Fig. 8): An active sunspot group within about  $40^\circ$  of c.m.p. A radial magnetic field has been established; however, it does not at present contain the earth, which is still surrounded by the general disordered field.

Cosmic rays produced in a flare in such a sunspot group would not have an easy route of access to the earth defined for them by the magnetic fields in space. To reach the earth, the radiation would have to diffuse through the general disordered field. As a result of this diffusion, the time scale of the rising intensity phase would be greater than if direct access (case A) were possible. Furthermore, the fact that the earth is surrounded by scattering centers would cause the radiation to appear to come from all directions quite early during the flare effect. That is, the apparent source would be broader, and the radiation would appear to become isotropic more rapidly than in case A.

*Case C* (Fig. 9): An active sunspot group within  $40^\circ$  of c.m.p., and a large flare has occurred in the group a day or two previously. The cloud of matter, with its entrained disordered magnetic fields, surrounds the earth, and the terrestrial cosmic-ray intensity is consequently depressed by the early stages of a Forbush decrease.

If another flare in the sunspot group were to generate cosmic radiation, the radial magnetic





Case C of the proposed model for the magnetic fields in the sun-earth region.

would define an easy route for part of the  
ce to the earth. For the remainder of  
ourney to the earth, however, the particles  
have to diffuse through the strong, dis-  
ed fields associated with the Forbush de-  
e. This diffusion would delay the arrival of  
adiation at the earth and would lengthen  
me scale of the increasing intensity phase  
flare effect. The relatively strong field, and  
aotic nature, would result in a mean free  
for scattering which would be shorter than  
in the general disordered field. Conse-  
ly, the time required for the radiation to  
e to the vicinity of the earth and the time  
of the increasing intensity phase could be  
derably greater than in case B. Further-  
the radiation would appear to reach  
py very rapidly.

Table 2 the seven largest flare effects ob-  
d by ground-level detectors have been  
t, together with the position on the solar  
t which the flare occurred and the prevail-  
rrestrial cosmic-ray conditions at the time.  
he basis of these data, we have determined  
of the previously considered cases fits the  
l conditions most closely. For example, the  
1959 event occurred when the flare-produce-  
center was about 30° from c.m.p., and soon  
a Forbush decrease, and hence fitted case  
e wish to emphasize that the class to which  
the effect was assigned was determined solely  
e conditions existing in the solar system  
to the occurrence of the solar flare.

TABLE 2. Details of the Seven Largest Cosmic-Ray Increases

The time scales are those derived from the rise times, and are relative to that of the May 1960 event.

| Event         | Position of Flare on Solar Disk | Cosmic-Ray Conditions, days since previous Forbush decrease | As- signed Class | Time Scale |
|---------------|---------------------------------|---|------------------|------------|
| May 1960      | 90°W                            | 5   | A                | 1          |
| February 1956 | 80°W                            | 5   | A                | 4          |
| November 1949 | 70°W                            | ...   | A                | 6          |
| March 1942    | 90°W                            | 7   | A                | 7          |
| February 1942 | 4°E                             | ...   | B                | 13         |
| July 1946     | 15°E                            | ...   | B                | 17         |
| July 1959     | 30°W                            | 1½  | C                | 36         |

Comparing the last two columns of Table 2, assigned class and relative time scale, we find good agreement with the predictions, for it can be seen that the four events with the shortest rise times were assigned to class A, the two events with the next longest rise times to class B, and the event with the longest time scale to class C. In addition, it was shown in section 1 that im- pact zones were not observed during the July 1959 event, an observation in accord with the predictions for class C.<sup>3</sup>

The rise times of the seven largest flare effects are therefore in good agreement with the theo- retical model. Recently it has been shown [*Reid and Leinbach*, 1959] that radio techniques can be used to detect the arrival at the earth of cos-

<sup>3</sup> Note added in proof. For class A of our model, the shortest time scales would occur when the magnetic field contained few scattering centers. This would also result in the radiation arriving at the earth being well collimated along the mag- netic lines of force and, consequently, impact zones would be well defined. A recent analysis has shown that during the May 4, 1960 flare event, impact zones were very well defined, in good agree- ment with theory. However, the analysis has sug- gested that at the time of the May 4 event, the lines of force in the ordered field were not radial, but inclined to the sun-earth line by about 70° to the west in the plane of the ecliptic.

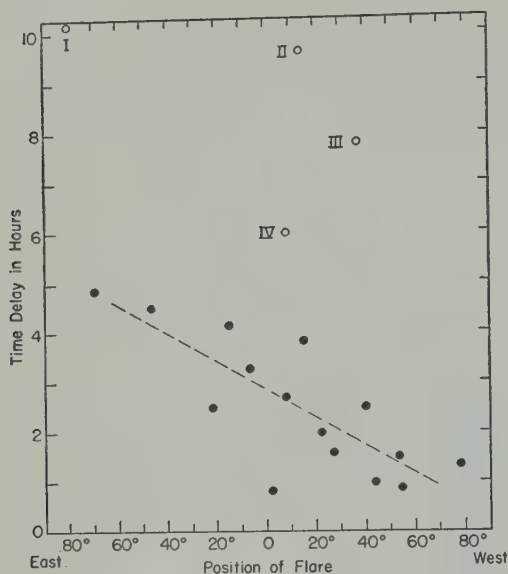


Fig. 10. The time delay between the visual observation of a flare and the start of the subsequent polar-cap blackout, plotted against the position of the flare on the solar disk. The four events represented by the open circles occurred on the following dates: I, March 23, 1958; II, February 10, 1958; III, September 2, 1957; IV, September 21, 1957.

mic rays produced in solar flares. This method reveals that a relatively large number of flares produce some cosmic radiation, and therefore a further test of the theoretical model is possible. To this end, Figure 10 shows the delay observed between the occurrence of a flare and the first indication of the arrival of cosmic radiation at the earth, plotted against the position on the solar disk at which the flare was observed (data derived from *Reid and Leinbach*, [1959]). It can be seen that the delays tend to be greatest for flares on the eastern half of the solar disk and least for flares near the western limb. With regard to the four events indicated by the open circles, their great deviation from the region occupied by the remainder of the points makes them somewhat suspect. The abnormally long delay of the September 2, 1957, event might possibly be due to the fact that it occurred when the terrestrial cosmic-ray intensity was greatly depressed by a Forbush decrease (case C), but no such explanation can be given for the other three anomalous events.

*Fifty per cent decay time.* It was demonstrated in section 3 that the rise and 50 per cent decay times vary from event to event by approximately the same factor. Clearly, two separate mechanisms, one to explain the rise time and the other to explain the decay time, would not, in general, predict such a behavior. The most natural explanation of the observed fact would be that a single feature of interplanetary space determines the time scales of both the rising and the decaying intensity phases of the flare effect.

Consider case A of the model we have been examining. The motion of cosmic-ray particles through a radial field in which there are irregularities approximates roughly to a random walk in which the motions perpendicular to the lines of force are eliminated by spiraling about the lines of force. In cases B and C the particle motion also approximates to a diffusion process. Simplifying the problem by considering it to be diffusion in a semi-infinite, one-dimensional medium [*Carslaw and Jaeger*, 1947], the cosmic-ray intensity at a distance  $x$  from the sun, at time  $t$  after injection, would be proportional to  $(kt)^{-1/2} \exp(-x^2/4kt)$ , where  $k$  is the diffusion coefficient in the direction of the lines of force. Since  $k$  and  $t$  only appear as the product  $kt$ , the shape of this intensity against time curve is unaffected by the value of  $k$ . But the rise and 50 per cent decay times are individually changed by the same factor: thus, if  $t_1$  is the time giving maximum intensity when  $k = k_1$  then, for  $k = k_2$ , the maximum intensity occurs at  $t_2 = k_1 t_1 / k_2$ .

The above very naive treatment, though predicting the observed correlation between changes in the rise and 50 per cent decay times, fails in that it predicts a  $t^{-1/2}$  dependence of intensity during the decreasing intensity phase, whereas a decay going as  $t^{-3/2}$  has been observed [*Meyer, Parker, and Simpson*, 1956]. A more rigorous treatment of the actual three-dimensional problem might remove this difficulty and consequently provide a good confirmation of the correctness of the proposed model.

*Concluding remarks.* Without invoking any models, certain properties of the solar flare effect have been demonstrated in this paper. Although the model we have considered may prove to be incorrect, these experimental facts apply

well founded and must consequently be based by any satisfactory theory. In summary the experimental facts are:

The time scale of the solar flare effect is a variable quantity, varying by a factor of at least as 36 from event to event.

The ratio of the rise time of a solar flare effect to the 50 per cent decay time is roughly constant. That is, the shape of the intensity vs. time curve is roughly the same for all effects.

The flares resulting in the four flare effects had the shortest time scales occurred on the eastern solar limb; those with the longest time scales occurred near the center of the solar disk. No impact zones were observed at the time of the flare effect that showed the greatest time

**Acknowledgments.** We are grateful to the many investigators who supplied data obtained during 1959. In particular, we desire to thank the Honorable Royal, Dr. H. Carmichael, Dr. A. G. E. Chilton, Dr. J. A. Lockwood, Dr. P. L. Marsden, Dr. P. Meyer, Dr. M. A. Pomerantz, Dr. D. C. Dr. A. E. Sandstrom, Prof. J. A. Simpson, Prof. B. Trumpy for the data presented in this paper. We should like to thank Prof. B. Rossi for his discussion during the course of this work. R. A. Almeida was supported by a fellowship from Conselho Nacional de Pesquisas (Brazil). K. G. McCracken was sponsored by the World Wide Research Scientists project of the U. S. Government International Co-operation Administration).

This work is supported in part through AEC Contract AT(30-1)-2098, by funds provided by the Atomic Energy Commission, the Office of Research, and the Air Force Office of Scientific Research; and in part, also, by a grant from the National Science Foundation.

#### REFERENCES

- D. K., and M. A., Pomerantz, The cosmic ray increase of 17 July, 1959, *Can. J. Phys.*, **38**, 322-333, 1960.
- Carmichael, H., and J. F. Steljes, Unusual cosmic ray fluctuations on July 17 and 18, 1959, *Phys. Letters*, **3**, 392-394, 1959.
- Law, H. S., and J. C. Jaeger, *Conduction of Heat in Solids*, Clarendon Press, Oxford, 1947.
- Chen, G., K. Greisen, P. Morrison, T. Gold, and Hayakawa, The cosmic ray flare effect, *Nuovo Cimento Suppl.*, **8**, 161-168, 1958.
- Chen, L. I., *Cosmic Ray Variations*, (translated by U. S. Technical Documents Liaison Office), The Publishing House for Technical and Theoretical Literature, Moscow, 1957.
- Firor, J., Cosmic ray intensity time variations, IV, Increases associated with solar flares, *Phys. Rev.*, **94**, 1017-1028, 1954.
- Fonger, W. H., Cosmic ray intensity time variations, II, Energy dependence of 27-day variation, *Phys. Rev.*, **91**, 351-361, 1953.
- Forbush, S. E., and B. F., Burke, Absorption of cosmic radio noise at 22.2 Mc/sec following a solar flare of February 23, 1956, *J. Geophys. Research*, **61**, 573-575, 1956.
- Forbush, S. E., T. B. Stinchcomb, and M. Schein, The extraordinary increase of cosmic ray intensity on November 19, 1949, *Phys. Rev.*, **79**, 501-504, 1950.
- Lüst, R., and J. A. Simpson, Initial stages in the propagation of cosmic radiation produced in solar flares, *Phys. Rev.*, **108**, 1563-1576, 1957.
- McCracken, K. G., The production of cosmic radiation by a solar flare on August 31, 1956, *Nuovo cimento*, **13**, 1074-1080, 1959.
- McDonald, F. B., and W. R. Webber, Proton component of the primary cosmic radiation, *Phys. Rev.*, **116**, 194-205, 1959.
- Meyer, P., E. N. Parker, and J. A. Simpson, Solar cosmic rays of February 1956, and their propagation through interplanetary space, *Phys. Rev.*, **104**, 768-783, 1956.
- Morrison, P., Solar origin of cosmic ray time variations, *Phys. Rev.*, **101**, 1397-1404, 1956.
- National Bureau of Standards, Central Radio Propagation Laboratory, Boulder, Colorado, *Solar-Geophysical Data*, issued August 1959.
- National Bureau of Standards, Preliminary Report of Solar Activity, Boulder, Colorado, issued May 13, 1960.
- Palmeira, R. A. R., and R. W. Williams, Rapid decrease of cosmic ray intensity. *Nuovo cimento*, **8**, 352-355, 1958.
- Parker, E. N., Cosmic ray modulation by the solar wind, *Phys. Rev.*, **110**, 1445-1449, 1958.
- Piddington, J. H., Interplanetary magnetic field and its control of cosmic ray variations, *Phys. Rev.*, **112**, 589-596, 1958.
- Quenby, J. J., and W. R. Webber, Cosmic ray cut-off rigidities and the earth's magnetic field, *Phil. Mag.*, **4**, 90-113, 1959.
- Reid, G. C., and Leinbach, H., Low-energy cosmic ray events associated with solar flares, *J. Geophys. Research*, **64**, 1801-1803, 1959.
- Rose, D. C., K. B. Fenton, J. Katzman, and J. A. Simpson, Latitude effect of the cosmic ray nucleonic and meson components from the Arctic to the Antarctic, *Can. J. Phys.*, **34**, 968-984, 1956.
- Rose, D. C., and J. Katzman, The sudden increase in cosmic ray intensity of February 23, 1956, *Can. J. Phys.*, **34**, 884-887, 1956.
- Wilson, B. G., D. C. Rose, and M. A. Pomerantz, *Can. J. Phys.*, **38**, 328-331, 1960.

(Manuscript received June 15, 1960; revised July 7, 1960.)





## Observations of Solar Flare Radiation and Modulation Effects at Balloon Altitudes, July 1959

A. EHMERT, H. ERBE, AND G. PFOTZER

*Max-Planck Institut für Aeronomie  
Lindau a. Harz, Germany*

AND

C. D. ANGER AND R. R. BROWN

*Department of Physics  
University of California  
Berkeley, California*

**Abstract.** Observations of cosmic-ray time variations in the stratosphere were obtained over Berkeley, California (USA), and Lindau/Harz (Germany) during the period of outstanding solar activity in July 1959. Whereas over Berkeley (geomagnetic latitude  $44^{\circ}\text{N}$ ) only modulation of cosmic radiation by solar-controlled processes was observed, additional radiation accelerated in the vicinity of the sun (approximately 46 per cent above normal) was observed over Lindau (geomagnetic latitude  $52^{\circ}\text{N}$ ) on July 15. On the assumption that the excess radiation consisted almost entirely of protons, its energy spectrum had the form  $N(E)dE \sim E^{-4}$  and extended up to 400 Mev. According to conventional geomagnetic theory, individual protons of these energies are normally forbidden at geomagnetic latitude  $52^{\circ}\text{N}$ . Their presence over Lindau on July 15 indicates that they were admitted by some mechanism which temporarily lowers the cutoff rigidity at this latitude during a magnetic storm. There is also some indication of further examples of excess radiation over Lindau at atmospheric depths down to 30 km Hg on July 13, 21, 23, and 31.

**Introduction.** During the IGY numerous stations at ground level recorded time variations of intensity of cosmic radiation. These observations were supplemented by many measurements using detectors carried aloft by balloons and satellites. The observations now available show that the time variations of cosmic radiation are controlled almost exclusively by solar activity. The following types of time variations of radiation intensity are distinguished: (a) those generally understood as a solar modulation of cosmic radiation, and (b) those resulting from solar flare radiation superimposed on the modulated level of cosmic radiation. Variations of the first type are always observed from the top of the atmosphere to ground level and from the poles to the equator (Forster, 1959). Secular variations with the cycle of solar activity). In time variations of the second type, radiation of solar origin, mainly with energies comparable to those found in cosmic radiation, follows promptly after some

outstanding solar flares and is observed at ground level and low latitudes. Only 6 events of such magnitude have been observed between 1942 and the present. More frequent examples of flare radiation are found at balloon altitudes for latitudes exceeding  $50^{\circ}$ . These events do not produce cosmic-ray effects at ground level, because of lower average energy of the radiation.

The second type of intensity variation is not as well understood as the first, owing to the much smaller numbers of simultaneous observations at high altitudes and different locations. The investigation described in the present paper had the purpose of studying time variations of cosmic radiation by means of simultaneous or at least overlapping balloon flights at opposite sides of the earth. The flights were planned as a cooperative program between the cosmic-ray groups of the University of California and the Max-Planck Institut für Aeronomie. Preliminary results [Brown and D'Arcy, 1959] of other balloon flights carried out in Alaska by the Univer-

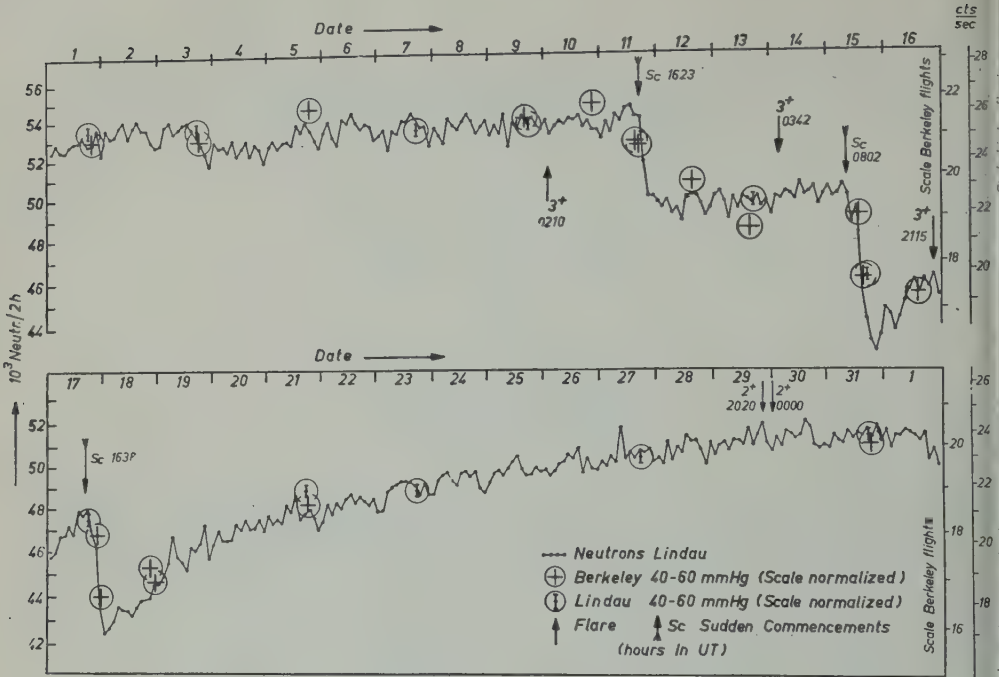


Fig. 1. Time variations of cosmic radiation during July 1959, measured at balloon altitudes and ground level.

sity of California group during the same time interval have been published elsewhere.

2. *Instrumentation.* The cosmic-ray-intensity measurements at balloon altitudes were obtained at both stations using single Geiger counters (Victoreen type 1B85). The pulses from the counters were fed to scalers; the switching of the end stages of the scalers controlled subcarrier frequencies that modulated the transmitter. In a similar manner the atmospheric pressure was telemetered to the ground, subcarrier frequencies being controlled by a baroswitch. All circuitry and counters were mounted in a styrofoam box. The absorber surrounding the counter volume amounted to 140 mg/cm<sup>2</sup> of styrofoam and 30 mg/cm<sup>2</sup> of aluminum. It was found that the counting rates of individual counters resulting from irradiation by a 1 Mc  $\gamma$ -ray source at a distance of 1 meter differed by about  $\pm 1$  per cent at an average rate of 40 counts/sec. Losses resulting from counter deadtime (150  $\mu$ sec) did not exceed 0.5 per cent at a counting rate of 20 to 25 counts/sec, similar to that found at high altitudes. In view of this, correc-

tions in comparing the results of the various flights can be neglected.

3. *Results.* The scheduling of coordinated balloon flights in July 1959 was very fortunate because three solar flares of importance 3+ occurred on July 10, 14, and 16. These flares gave rise to intense proton bombardment of the polar cap regions, resulting in strong absorption of cosmic radio noise [Leinbach and Reid, 1959] and strong attenuation of signals on ionospheric scatter circuits in the polar regions [Brown, 1959]. In addition, intense solar corpuscular streams gave rise to strong aurora as well as intense magnetic storms after each of the flares.

Cosmic radiation time variations resulting from the solar flare outbursts observed well within the atmosphere at Berkeley and Lindau are summarized in Figure 1. The small dots connected by solid lines represent the bihourly counting rates of the neutron monitor at Lindau plotted on a logarithmic scale versus time. Crosses and vertical bars surrounded by circles indicate the time of balloon flights and the corresponding omnidirectional intensities observed.



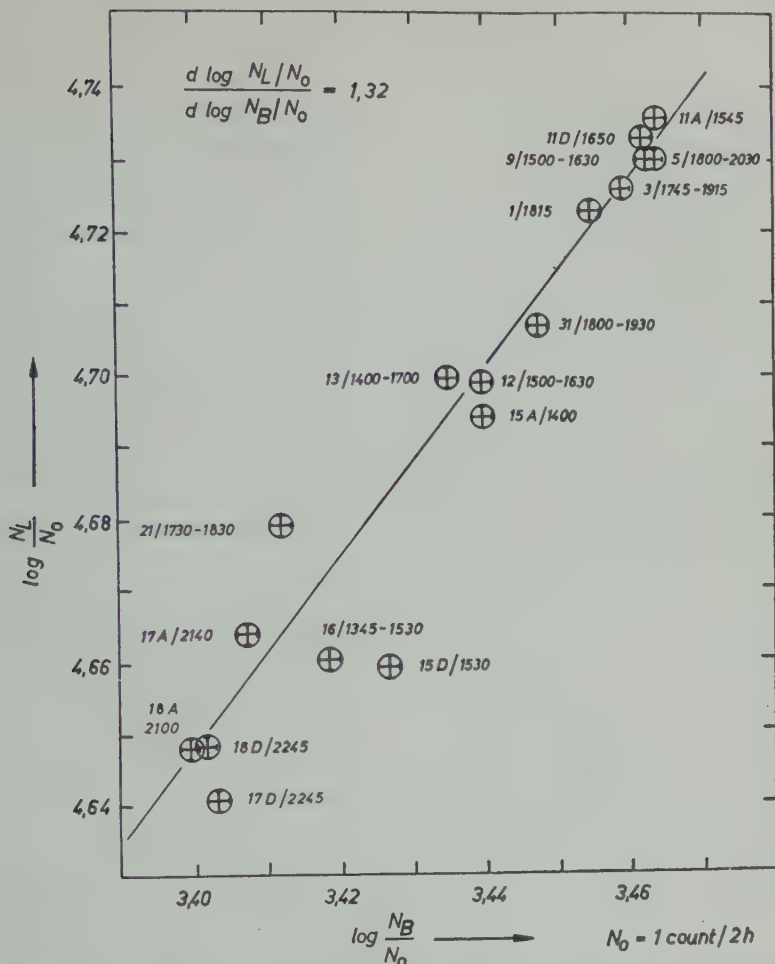


Fig. 2. Comparison of Lindau and Berkeley neutron monitor counting rates (logarithmic scales).

in the stratosphere between pressure levels 60 and 40 mm Hg. These values correspond to the counting rates at the maximum of the intensity versus pressure curve of cosmic radiation that appears under normal circumstances.

To make the flight data fit closely with the neutron data, different bases for the corresponding logarithmic scales have been used. Their proportions are 1:0.8:0.59 for the Lindau neutron monitor and the Berkeley and Lindau flight data, respectively. The physical meaning of this normalization will be discussed in detail in section 4.

Instead of the Lindau neutron monitor data, the Berkeley neutron monitor data could have been used as a reference. This follows im-

mediately from Figure 2, where the counting rate of the Berkeley monitor is plotted against that of the Lindau monitor, both on logarithmic scales, for the times when balloon flights were made at Berkeley. From this, it is found that the neutron monitor variations are related according to

$$[d(\ln N_{\text{Lindau}})]/[d(\ln N_{\text{Berkeley}})] = 1.32$$

Summarizing the results shown in Figures 1 and 2, it is to be noted that the intensity variations

$$\delta J(t)/J(t) = d(\ln J(t))$$

of cosmic radiation measured well within the atmosphere at different locations on the earth

are proportional to one another. This holds, as will be seen below, in July from ground level up to atmospheric depths of approximately 40 mm Hg. The only exception during this period occurred on July 15, following the onset of a sudden-commencement magnetic storm.

At atmospheric depths less than 40 mm Hg, the flights over Berkeley did not show any pronounced deviations from the expected values based on the correlation with neutron data. Therefore, further discussion of the balloon flights will be confined to the Lindau flights, in which significant deviations from normal were observed.

In Figure 3 the counting rates of the Geiger counters measured over Lindau in July are plotted versus the logarithm of the atmospheric pressure, in units of millimeters of mercury. Time reference points, in UT, are at the top and bottom of the figure, referring to balloon ascent and descent, respectively. The time interval between two points amounts to 2 minutes; the values themselves have been averaged over 4 minutes. The smooth curves represent the expected altitude variation of cosmic radiation as discussed in sections 4 and 5. Since the instruments were carried aloft by only one balloon on the Lindau flights and descended rapidly by parachute, greatest weight must be given to measurements obtained during the balloon ascent.

On July 1, 3, 7, and 9 the intensity versus pressure curves were practically identical. After the solar flare of importance 3+ on July 10, the neutron counting rate at Lindau showed a Forbush decrease of 11 per cent, remaining close to that level until July 15 (Fig. 1). As would be expected, the counting rate at high altitudes on July 13 was somewhat lower than before the Forbush decrease. However, comparison with the counting rate-pressure curve based on neutron monitor observations at ground level shows that the decrease at balloon altitudes was smaller than expected if the counter responded only to cosmic radiation normally present at this latitude. On July 15, when the counting rate was still 8 per cent below the level before the Forbush decrease on July 11, a second Forbush decrease, amounting to 15 per cent, was observed with the neutron monitor. The neutron monitor rate dropped now to a level 21 per cent below

that observed before July 11. The decreased counting rate at the 50 mm Hg pressure level, however, was not considerably smaller than expected. Furthermore, instead of going through a maximum at this level in the atmosphere, the counting rate remained relatively flat and then increased with decreasing atmospheric pressure, reaching a level 46 per cent above normal at the 10 mm Hg pressure level. This observation, representing a clear case of excess radiation of solar origin, will be discussed in more detail in section 5.

After some recovery of the radiation level from that reached on July 15, a third Forbush decrease occurred on July 18 and the counting rate of the Lindau neutron monitor dropped to a level 23 per cent below that on July 11, the lowest level that has been observed up to the present time. After the third Forbush decrease, balloon flights on July 21 and 23 gave counting rates at depths above the intensity maximum which were somewhat in excess of that to be expected on the basis of the neutron monitor counting rate at ground level.

4. *Normalization of data.* The normalization data, mentioned briefly in sections 2 and 3, is based on numerous measurements [Pomeroy and Agarwal, and Potnis, 1958; Winckler and Peterson, 1958; Brown, 1959; Henkel, Lockwood, and Trainor, 1959; Ehmert and Erbe, 1959] during the IGY in which relative time variations of cosmic-ray intensity at balloon altitudes were compared with the time variations of neutron monitor counting rates on the ground. Erbe [1959], for example, found that the  $J_h(t)$  versus  $\ln N(t)$  plot can be represented to good approximation by a straight line, where  $J_h(t)$  is the vertical flux of particles at time  $t$  and altitude  $h$ , and  $N(t)$  is the neutron counting rate at ground level. This result is equivalent to

$$[d(\ln J_h(t))]/[d(\ln N(t))] = S_h$$

where  $S_h$  is independent of time but increases with increasing altitude. Integrating this equation gives

$$J_h(t) = J_h(t_0)(N(t)/N(t_0))^{S_h}$$

It is assumed now that this relation is applicable in dealing with cosmic radiation which is modulated in the planetary system by a solar mechanism, as yet unknown. If this is so, the level  $J_h$

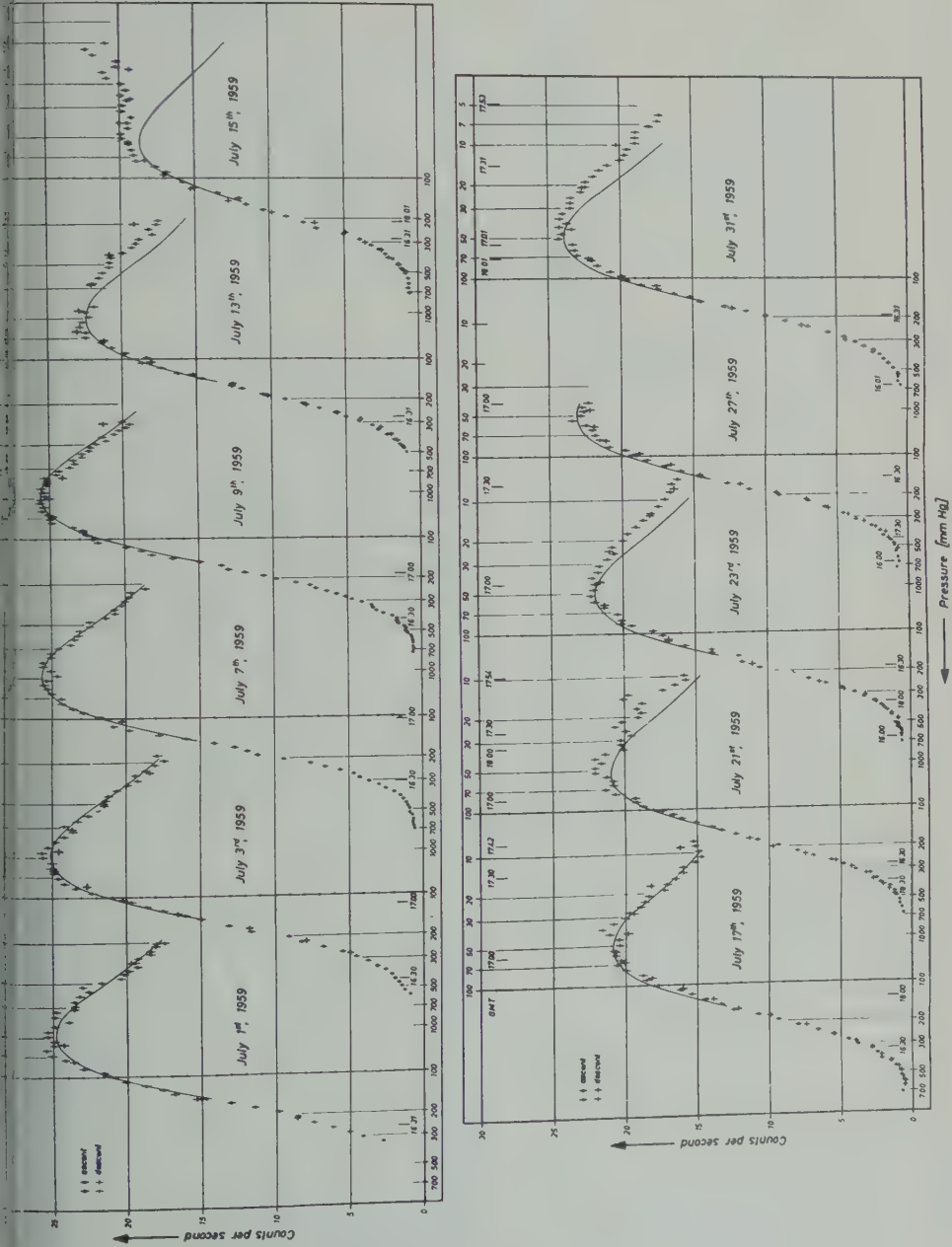


Fig. 3. Intensity-pressure curves measured over Lindau in July 1959 (smooth curves represent the expected variation derived from ground-level neutron monitor data).



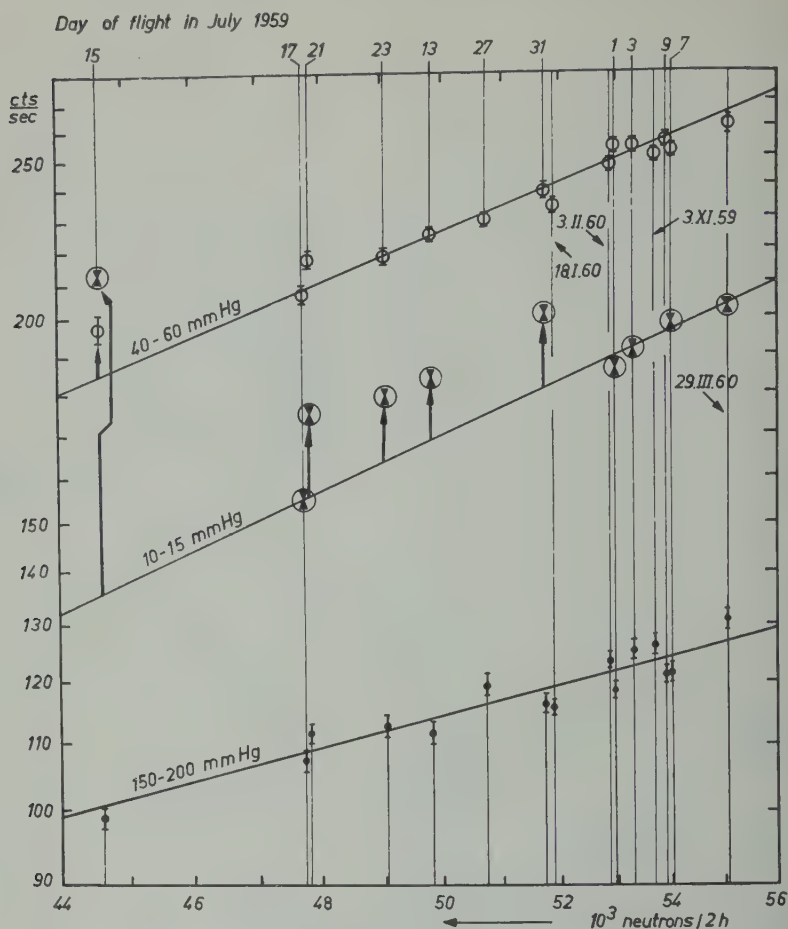


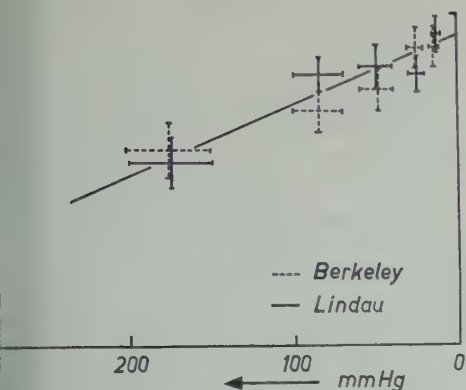
Fig. 4. Counting rate of single Geiger counters at different atmospheric depths versus counting rate of ground-level neutron monitor (Lindau).

of cosmic radiation at time  $t$  at altitude  $h$  can be calculated from the level  $J_h(t_0)$  at time  $t_0$  and the counting rates  $N(t)$  and  $N(t_0)$  of a ground-level neutron monitor at times  $t$  and  $t_0$ .

The exponent  $S_h$  has been determined at different altitudes by Erbe for the vertical flux of cosmic radiation. Since the exponent cannot be expected to be the same for radiation measured with an omnidirectional detector, an exponent  $S_h$  was derived for this type of detector from the balloon flights and neutron rates at Lindau, as shown in Figure 4.

The values of  $J_h(t)$  for atmospheric pressures in the ranges 150 to 200 and 40 to 60 mm Hg correlate well with  $N(t)$ , as shown by the grouping of the points about straight lines in Figure

4. At these depths in the atmosphere,  $S_h$  is determined unambiguously. This holds also for the flights over Berkeley at all pressures. However, at the highest altitudes reached during Lindau flights, 10 to 15 mm Hg, marked deviations from a straight line are found, as in Figure 4. In order to examine them more closely, it is assumed that before the first sharp decrease on July 11 only cosmic radiation of galactic origin was present (July 1, 3, and 7). From Erbe's results it is known that  $S_h$  increases with increasing altitude. Thus, the line representing the variation of cosmic radiation at this depth must be steeper than for intermediate higher atmospheric pressure. A lower line for  $S_h$  at this depth is obtained from the slope



5. Variation of factor  $S_h$  with atmospheric depth.

connecting the points for July 1, 3, and 7 that on July 17. If this procedure is accepted, it is seen that excess radiation was present at small atmospheric depths on July 21, 23, and 31 as well as on July 15. The length of the arrows in Figure 4 connecting the 15 mm Hg. line and the observed points represents the extent to which the intensity of total radiation exceeded normal cosmic radiation on these occasions. Similar conclusions also for the pressure interval 20–30 mm Hg. For the sake of clarity, the data for this event were not included in Figure 4.

The exponents of  $S_h$  for flights over Lindau and Berkeley, determined by the procedure outlined above, are shown in Figure 5. The differences between exponents  $S_h$  for these two locations are within the limits of error and do not appear to be significant.

*The event of July 15.* To determine the amount of excess radiation over Lindau on July 15 the normal cosmic-ray intensity versus pressure curve for this day was calculated using neutron and neutron data for July 17 as a reference. Thus, it follows that

$$J_h(15) = J_h(17)(N(15)/N(17))^{S_h}$$

The flight on July 17 was chosen as a reference because the neutron counting rate  $N(17)$  was higher than for any other day where neutron data were available. It is evident that for  $N(15)/N(17) \approx 1$  the error in the exponent  $S_h$  does not influence the value of  $J_h$  appreciably. Thus, an error of  $\pm 17$  per cent in  $S_h$  implies an

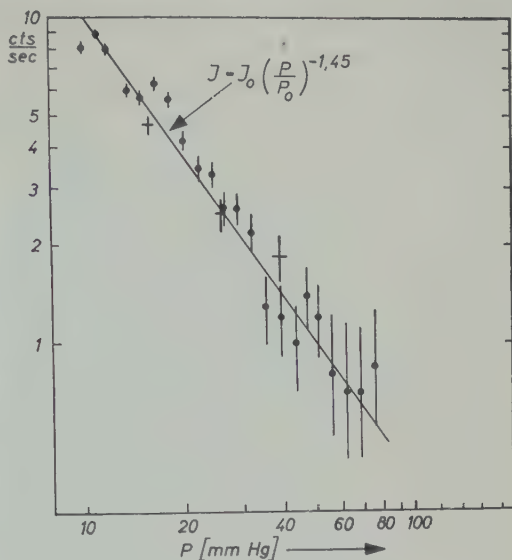


Fig. 6. Excess radiation on July 15, 1959 (Lindau).

error of  $\pm 1$  per cent in  $J_h(15)$  for  $N(15)/N(17) = 1.07$ . Subtracting the normalized counting rates represented by the smooth curve in Figure 3 from the counting rates actually measured (full dots) gives the intensity of the excess radiation as a function of atmospheric depth, as shown in Figure 6.

There is now evidence from emulsion exposures [Ney, Winckler, and Freier, 1959; Winckler, 1959] at balloon altitudes that the radiation coming from solar flare outbursts consists mainly of protons. It is assumed that this was also true on July 15 over Lindau. This assumption, together with the assumption that the excess radiation did not vary appreciably with time during the ascent and descent of the flight on July 15, leads to an estimate of the energy spectrum of the excess radiation. Thus, from Figure 6, the omnidirectional flux  $J'$  of the excess radiation may be represented to a good approximation by

$$J' = C'p^{-1.45}$$

where  $p$  is the atmospheric pressure and  $C'$  is a constant. Taking the energy spectrum to be a power law, a Gross transformation leads to the same form for the vertical flux  $J$

$$J = CR^{-1.45}$$

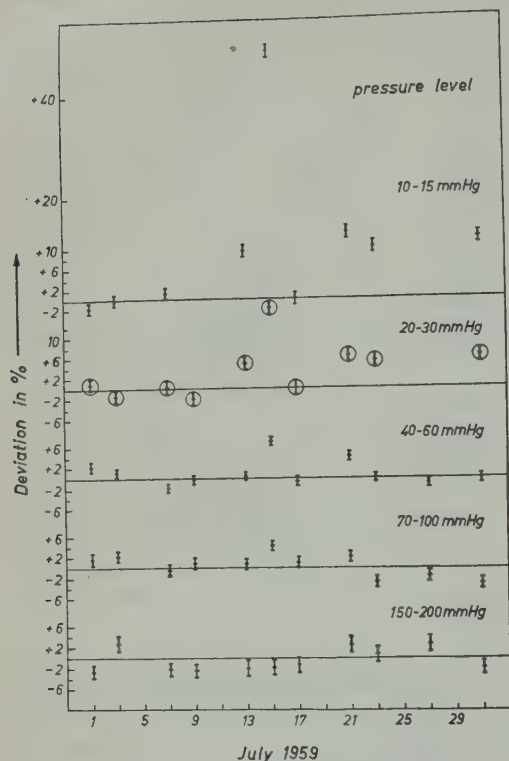


Fig. 7. Deviation of cosmic-ray intensity from the expected value based on ground-level neutron monitor data (Lindau).

where  $R$  is now the range of the proton in air. Denoting the kinetic energy by  $E$  and the differential energy spectrum of the protons by  $f(E)dE$ , it follows that

$$f(E) dE = (dJ/dR)(dR/dE) dE$$

Using the range-energy relationship [Wilson, 1947]

$$R \sim E^{1.8}$$

gives

$$f(E) dE \approx R^{-2.45} E^{0.8} dE$$

or

$$f(E) dE \sim E^{-3.6} dE$$

for the energy spectrum of the protons.

Since the excess radiation was observed to atmospheric depths of 70 mm Hg, the spectrum extends up to energies of about 400 Mev. According to geomagnetic theory [Quenby and

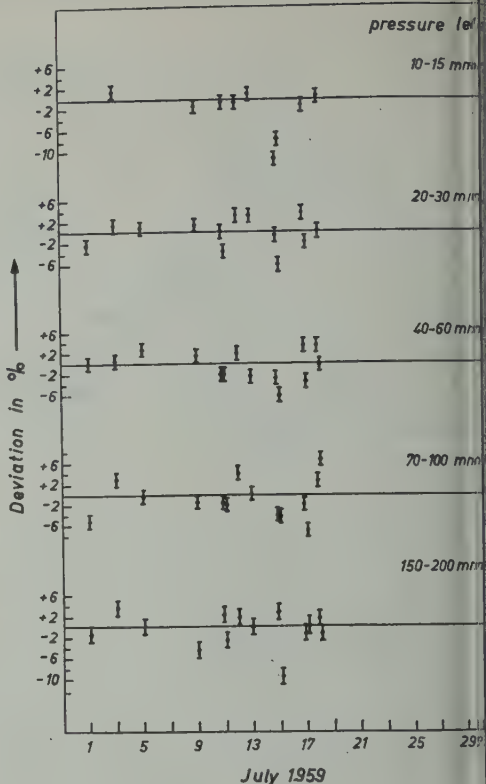


Fig. 8. Deviation of cosmic-ray intensity from the expected value based on ground-level neutron monitor data (Berkeley).

Webber, 1959] only protons with kinetic energy greater than 1.65 beV are allowed to enter the atmosphere at Lindau. From this discrepancy it appears that a temporary failure of the magnetic cutoff occurs during some phase of intense geomagnetic storms, admitting high energy particles to regions normally forbidden to them. Similar conclusions have been reached by the University of Minnesota group [Winckler, 1959; Ney, 1959] from an analysis of proton tracks in nuclear emulsions exposed on July 15 at high altitudes over Minneapolis and by the University of Chicago group [Fenton and Sisson, 1959] from vertical flux measurements with counters at balloon altitudes over Chicago on July 15. Similar instances have been reported by the Minnesota group [Ney, Winckler, Freier, 1959], protons well below geomagnetic cutoff energies arriving over Minneapolis on March 1958 and May 1959.



*Flight summary.* Aside from the outstanding event of July 15, smaller deviations from expected intensities of cosmic radiation are indicated in Figure 2 for other flights over Lindau. The amounts of these deviations are given for different days and pressure levels in Figure 7, and for comparison the same plot is given for the Berkeley flights in Figure 8. It is seen that the points for Berkeley scatter around the expected values, showing no marked deviations such as were observed over Lindau on July 15. Therefore, it is concluded that no excess radiation reached the latitude of Berkeley ( $44^\circ\text{N}$  magnetic latitude, cutoff energy  $E_c = 3.6$  bev). For Lindau, it is seen that no excess radiation is present at atmospheric depths greater than 1000 mm Hg. For smaller atmospheric depths, starting from the event of July 15, excess radiation seems to have been present on July 13, 21, 23, and 31. On July 13, 21, and 23, it would seem reasonable to assume that there still remains the influence of the processes following the 3+ event on July 10 and 16 which gave rise to the Forbush decreases on July 11 and 18. In the period following the Forbush decrease of July 10, however, only a smooth recovery of the neutron level is observed on the ground, together with small but distinct diurnal variations between July 21 and 26. Magnetic disturbances do not reach any high level nor did ionospheric characteristics show any unusual features during this period. The only solar events of any significance were four flares on July 29, two of importance 2 at 0709 and 1213 UT and two of importance 2+ at 2020 and 2400 UT, all in the active region around  $11^\circ\text{N}$ ,  $37^\circ$  to  $34^\circ$  E. At present, it is difficult to find any other cause for the deviation of cosmic-ray intensity from the level expected on July 31 from neutron monitoring correlations; it is hoped that cosmic-ray observations made elsewhere in the same period will clarify this point.

*Discussion.* From the observations over Lindau on July 15, as well as others obtained by the Minnesota and Chicago groups at about the same time, it is evident that large temporary decreases of the geomagnetic cutoff rigidity occur during strong geomagnetic storms. Roth [1959] has suggested that such cutoff failures result from solar plasma distorting the magnetic field, the mag-

nitude of the changes in the cutoff rigidity depending on the depth of penetration of the plasma into the geomagnetic field, but the changes are limited mainly to high-latitude regions.

Such a temporary lowering of the cutoff rigidity would be difficult to detect in polar latitudes, where it is most pronounced, with ground-level neutron monitors, because of the rapid fall-off of the neutron yield function at low primary rigidities. At lower latitudes, such changes in the cutoff rigidity would be much smaller and any resulting effects probably hidden by the modulation of cosmic radiation by solar-controlled processes [Ehmert, 1960; Fan, Meyer, and Simpson, 1960].

At balloon altitudes, any changes in the cutoff rigidity would be more apparent, the intensity-depth curve at a given latitude taking on the form appropriate to a higher magnetic latitude. Such might be the case for the observations over Lindau on July 13, 21, 23, and 31, although the connection with magnetic activity is not entirely clear.

In conclusion, it should be pointed out that solar proton events, when observed at latitudes where low-energy radiation is normally forbidden, should be corrected for effects resulting from the cutoff failure before energy spectra are derived. Thus, the excess radiation over Lindau on July 15 may contain a mixture of low-energy solar protons and cosmic-ray protons normally forbidden at this latitude. The cosmic-ray protons, characterized by a slowly varying spectrum, have the effect of giving the excess radiation a spectrum which is flatter than the solar proton spectrum.

*Acknowledgments.* The research was supported in part by the Deutsche Forschungsgemeinschaft Bad Godesberg and in part by the Joint Program of the U. S. Office of Naval Research and the U. S. Atomic Energy Commission.

#### REFERENCES

- Bailey, D. K., Survey of solar proton events from VHF ionospheric scatter observations, Midwest Cosmic Ray Conference, 1959.
- Brown, R. R., Excess radiation at the Pfitzner maximum during geophysical disturbances, *J. Geophys. Research*, **64**, 323-329, 1959.
- Brown, R. R., and R. G. D'Arcy, Observations of solar flare radiation at high latitude during the

- period July 10-17, 1959, *Phys. Rev. Letters*, **3**, 390, 1959.
- Ehmert, A., On the modulation of primary cosmic ray spectrum by solar activity, *Proc. Moscow Cosmic Ray Conf. IV*, 142, 1960.
- Ehmert, A., and H. Erbe, Balloon measurements of time variations of primary cosmic radiation, *Proc. Moscow Cosmic Ray Conf. IV*, 80, 1960.
- Erbe, H., *Mitt. Max-Planck-Institut für Aeronomie*, **2**, 1959, Dissertation, Stuttgart, 1959.
- Fan, C. Y., P. Meyer, and J. A. Simpson, Cosmic radiation intensity decreases observed at the earth and in the nearby planetary medium, *Phys. Rev. Letters*, **4**, 421, 1960.
- Fenton, K. B., and J. A. Simpson, 200 Mev particles arriving over Chicago on July 15, 1959, Midwest Cosmic Ray Conference, 1959.
- Freier, P. S., E. P. Ney, and J. R. Winckler, Balloon observations of solar cosmic rays on March 26, 1958, *J. Geophys. Research*, **64**, 685-686, 1959.
- Henkel, J. E., J. A. Lockwood, and J. H. Trainor, Comparison of the cosmic-ray intensity at high altitude with the nucleonic component at ground elevation, *J. Geophys. Research*, **64**, 1427-1438, 1959.
- Leinbach, H., and G. C. Reid, Observations and interpretation of the high latitude absorption events of July 1959, Midwest Cosmic Ray Conference, 1959.
- Ney, E. P., Emulsion observations of solar protons in July, 1959, Midwest Cosmic Ray Conference, 1959.
- Ney, E. P., J. R. Winckler, and P. S. Freier, Protons from the sun on May 12, 1959, *Phys. Rev. Letters*, **3**, 183, 1959.
- Pomerantz, M. A., S. P. Agarval, and V. R. P. Direct observation of periodic variations of primary cosmic ray intensity, *Phys. Rev.* **109**, 1958.
- Quenby, J. J., and W. R. Webber, Cosmic ray cutoff rigidities and the earth's magnetic field, *Phil. Mag.*, **4**, 90, 1959.
- Rothwell, P., Magnetic cutoff rigidities of charged particles in the earth's field at times of magnetic storms, *J. Geophys. Research*, **64**, 2028, 1959.
- Wilson, R. R., Range, straggling and multiple scattering of fast protons, *Phys. Rev.* **71**, 1947.
- Winckler, J. R., Preliminary report on the interplanetary low energy solar cosmic ray outbursts occurring in July 1959, Preprint, University of Minnesota School of Physics, July 1959.
- Winckler, J. R., and L. Peterson, A large cosmic ray decrease accompanying the solar maximum of 1957, *Nature*, **181**, 1317, 1958.

(Manuscript received June 20, 1960.)

# Development of Multiple Radiation Zones on October 18, 1959

B. J. O'BRIEN AND GEORGE H. LUDWIG

*Department of Physics and Astronomy  
State University of Iowa  
Iowa City, Iowa*

**Abstract.** Data from two Geiger counters on Explorer VII indicate that anomalous radiation zones developed around 1000-km altitude on the low-latitude side of the outer radiation zone during a moderate magnetic storm on October 18, 1959. The zones were first detected as quite separate regions of increased counting rates, each region about  $0.5^\circ$  thick in latitude. Later, other narrow zones were seen to be superimposed on a  $6^\circ$ -thick zone of steadily increasing intensity. At one time, this broad zone contained radiation even more intense than that in the normal outer zone. The anomalous radiation disappeared within a few days as the outer zone returned to its prestorm condition. Correlation with ionospheric data indicates that much of the detected radiation was trapped for some hours at least. A qualitative explanation of the phenomena is suggested.

On several occasions two Geiger counters on satellite Explorer VII have measured multiple zones of high-intensity radiation at altitudes between 560 and 1100 km. In general, these zones have been detected on the low-latitude side of the outer radiation zone. The morphology of the development of these narrow multiple zones into a broad zone is studied for a period when a moderate magnetic storm occurred on October 18, 1959.

Explorer VII was launched on October 13, 1959. It is spin-stabilized against tumble, and the Geiger counters have uniform shielding over about  $3\pi$  steradians. The apparatus has been described in detail [Ludwig and Whelpley, 1960]. The Anton 302 counter and the Anton 112 counter are shielded by at least  $0.7 \text{ g cm}^{-2}$  of stainless steel and  $1.6 \text{ g cm}^{-2}$  of lead, respectively. They will efficiently count protons above about 20 and 30 Mev and electrons above about 1 and 3 Mev, and, less efficiently, the bremsstrahlung from electrons above about 30 and 80 keV, respectively.

A moderate magnetic storm began about 0100 UT on October 18, 1959, and lasted until about 0100 Z on October 19. The values of  $K_p$  over this period [National Bureau of Standards, 1959] and the general characteristics of the outer radiation zone are listed in Table 1. Northern-hemisphere observations were at alti-

tudes around 600 km over North America; those near Australia were around 1000 km.

In Table 1 the outer zone has been considered in three rather arbitrary intervals of  $\lambda_{dip}$ , where  $\lambda_{dip}$  is such that

$$2 \tan \lambda_{dip} = \tan \delta$$

where  $\delta$  is the angle of dip of the magnetic field line through the satellite. We define  $R$  as the ratio of 302 counting rate to 112 counting rate at the latitude  $\lambda'_{dip}$ , where the 302 rate is a maximum. We have considered the normal outer zone to be of a form similar to that seen on October 16 and 17, 1959, as shown in Figure 1, where the satellite was moving southward out of the zone. From consideration of data from Explorer IV [Van Allen, 1959], we believe that most of the counts in the outer zone probably are caused by bremsstrahlung of electrons of order 100 keV.

In Figure 2 we show the counting rates around 1531 Z on October 18, when the 112 counter with a scaling factor of 128 showed four narrow zones of high radiation intensity on the low-latitude side of the outer radiation zone over North America. As is indicated in Table 1, this structure was also present at 1350 Z and 1715 Z. As is seen in Figure 3, the structure on the latter pass was superimposed on an increased 'background' counting rate at low latitudes.



TABLE 1. Characteristics of the Outer Radiation Zone October 17-20, 1959

| Time      | $K_p$          | $\lambda_{dip} \gtrsim 63^\circ$ | $63^\circ > \lambda_{dip} \gtrsim 60^\circ$ |           | $60^\circ > \lambda_{dip}$   |
|-----------|----------------|----------------------------------|---|-----------|--|
|           |                |                                  | $\lambda'_{dip}$                            | $R$       |  |
| 17/1323 Z | 4 <sup>-</sup> | Normal                           | -62°  | 13        | Normal   |
| 17/2113 Z | 4 <sup>-</sup> | Normal                           | +62°  | 13        | Normal (Fig. 1)  |
| 18/1259 Z | 4 <sup>+</sup> | Reduced rates                    | -62°  | 13        | Normal   |
| 18/1350 Z | 5 <sup>-</sup> | Reduced rates                    | No data                                     |           | Twofold increase in 20-sec wide peak in 112 at 55°   |
| 18/1532 Z | 5 <sup>-</sup> | Reduced rates                    | +61°  | 7         | Four short increases in 112 around 58° (Fig. 2)  |
| 18/1715 Z | 4 <sup>-</sup> | Reduced rates                    | +62°  | 12        | Several 112 peaks around 56°-60°, with intense 20-sec-wide peak at 56°; small 30-sec-wide 302 peak around 58° (Fig. 3) |
| 18/1820 Z | 4 <sup>-</sup> | No data                          | -63°  | Uncertain | Broad and intense 112 and 302 peaks around -59°  |
| 18/1907 Z | 4 <sup>-</sup> | No data                          | ~62°  | ~8        | No data  |
| 18/2052 Z | 4 <sup>+</sup> | No data                          | +61°  | 13        | Intense 112 enhancement around 56° and 302 enhancement around 58° (Fig. 1)   |
| 19/1509 Z | 1 <sup>+</sup> | No data                          | +61°  | 15        | General enhancement at low-latitude edge (Fig. 2)  |
| 19/2029 Z | 3 <sup>+</sup> | No data                          | +60°  | 13        | Twofold increase 10 sec wide in 112 at 56°; small extension of low-latitude edge (Fig. 1)                              |
| 20/1214 Z | 1 <sup>+</sup> | Returning to normal              | -62°  | 15        | Normal   |
| 20/2006 Z | 0 <sup>+</sup> | Returning to normal              | +62°  | 13        | Normal (Fig. 1)  |

By 1820 Z a broad zone of enhanced radiation intensity was detected around  $\lambda_{dip}$  of -59°, south of Australia. A somewhat similar but more intense broad zone was then detected around 2052 Z over North America (Fig. 1). The intensity of this radiation gradually lessened, so that at 1510 Z on October 19 the additional radiation merely extended the broad outer zone to latitudes somewhat lower than normal (Fig. 2). By 2030 Z this effect was even smaller, although one small 112 peak was still present (Fig. 1). By 2007 Z on the next day, the low-latitude side of the outer zone appeared normal, i.e., similar to the prestorm condition (Fig. 1).

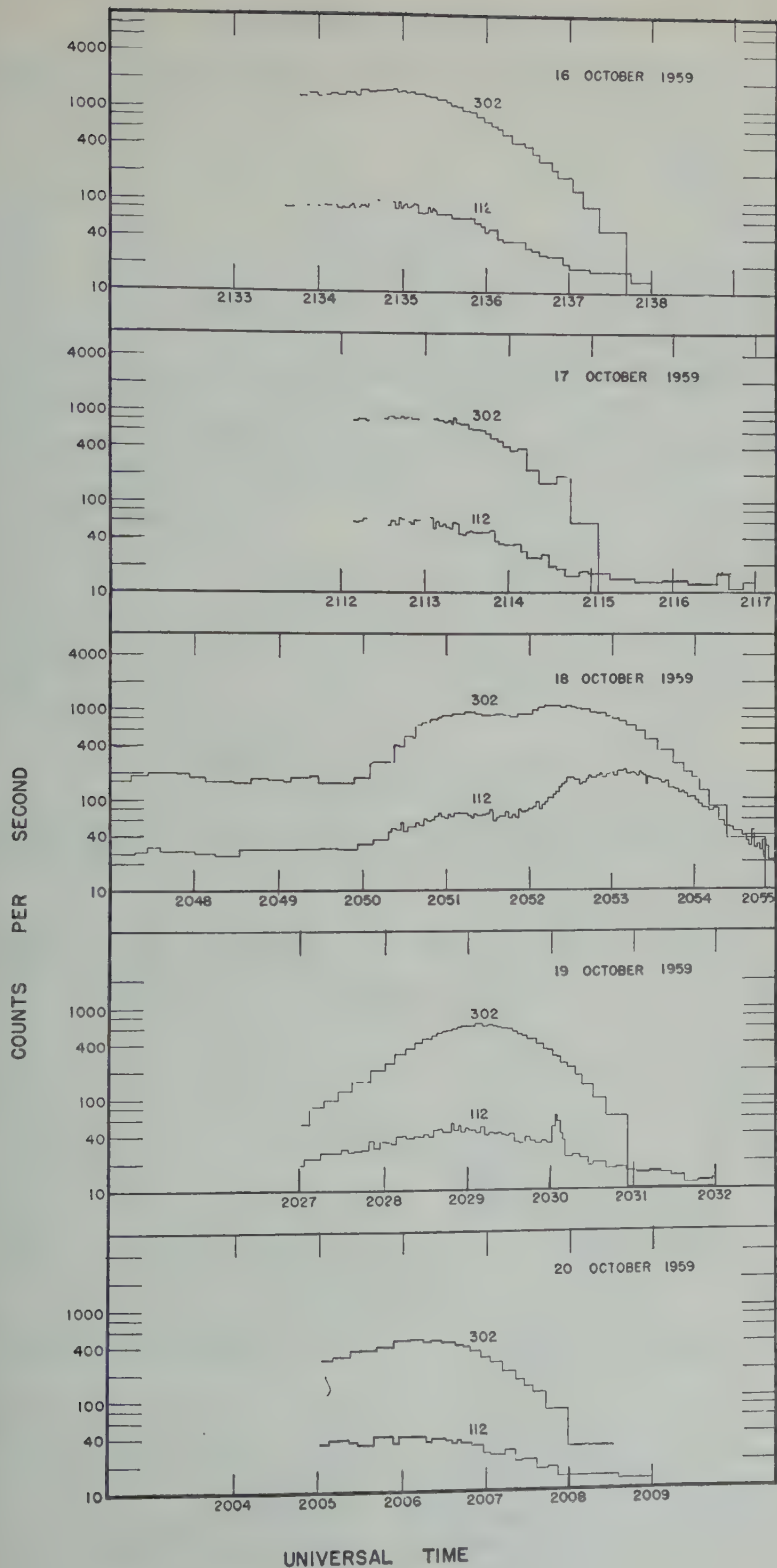
As can be seen from Table 1, the high-latitude (above 63°) trapped radiation was depleted early in the storm, and then gradually recovered over a period of days (see also Fig. 1). A similar course of events has been observed during several other magnetic storms since Explorer VII was launched.

From the data of Table 1, and by compari-

son of the two graphs in Figure 2, it appears that the storm also caused a temporary decrease in the intensity of the trapped electrons in the outer-zone peak of energies around 50 kev, those detected by the 302 counter but not the 112 counter.

From consideration of the satellite data alone it is not possible to determine the extent to which the anomalous radiation was trapped. We have studied the effects on the ionosphere over this period from stations geographically located near the regions where the anomalous zones were found (G. King, private communication, 1960), and we believe that most of the detected radiation was trapped for periods of hours at least.

At Fort Monmouth (-74.1°, 40.3°) the value of  $f_oF_2$  at 1530 Z was 9.8 Mc/s compared to the normal value of 11.5 Mc/s. At 2100 Z the value was normal at 13.5 Mc/s. Similar behavior occurred at Fort Belvoir (-77.1°, 38.8°). The satellite pass from 2052 to 2053 Z on



1. Counting rate of Geiger tubes 302 and 112 against time for similarly located passes of Explorer VII for October 16-20, 1959. The graphs are aligned by the time at which the satellite crossed geographical latitude  $40^{\circ}\text{N}$ , moving southward.

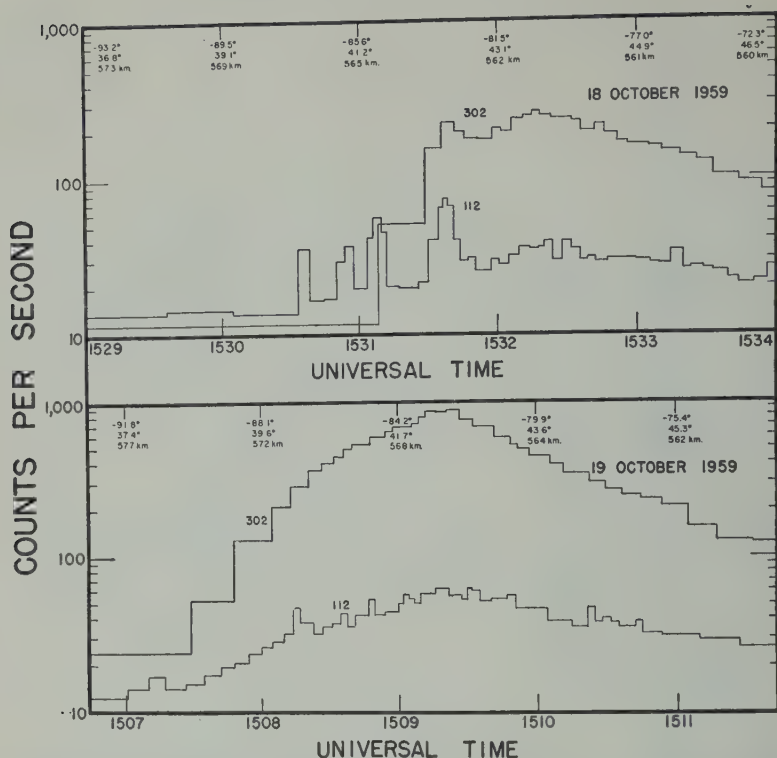


Fig. 2. Counting rate of Geiger tubes 302 and 112 against time for similarly located passes of Explorer VII on October 18-19, 1959. The graphs are aligned as in Figure 1, but the satellite is moving northward.

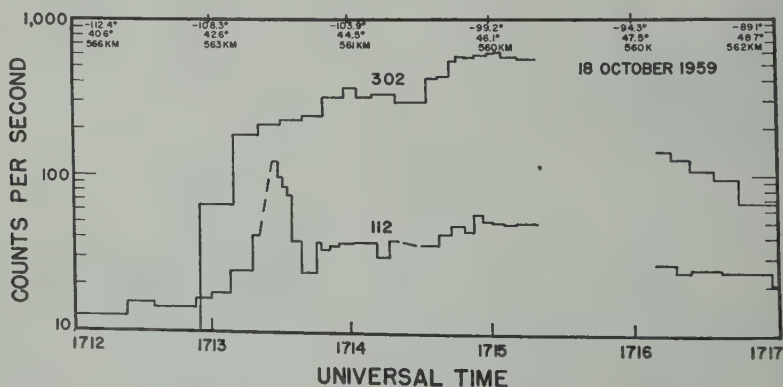


Fig. 3. Counting rate of Geiger tubes 302 and 112 against time for a pass around 1715 Z on October 18, 1959, as Explorer VII moved northward and passed through the outer zone.

October 18 was situated between these two stations. At White Sands ( $-106.5^\circ$ ,  $32.3^\circ$ ) the ionospheric disturbance lasted somewhat longer, with observed  $f_oF_2$  at 1530 and 2100 Z of 7.2 and 10.8 Mc/s compared with the normal values of 11.4 and 13.2 Mc/s. This amount of iono-

spheric disturbance is not considered Sporadic E ( $E_s$ ) activity at the station was low and showed no unusual features.

Although an association of the ionospheric and satellite data is necessarily tenuous, the evidence seems to indicate that much



alous radiation that was detected remained trapped for some hours at least, so that the energy flux fed by the radiation into the ionosphere was low at any instant.

The identification of the particles in the anomalous flux is also uncertain. It would appear most likely that they were electrons of somewhat higher average energy than those in the outer zone, whose bremsstrahlung produced the shielding gave the observed counting rate.

On this assumption the flux of electrons of about 30 kev which were detected at 30° Z on October 18 can be estimated as about  $10^{10}$  particles/cm<sup>2</sup> sec. Such a flux would certainly give severe ionospheric disturbances if the major portion were not trapped.

The essential features of this study may be summarized as follows. During a moderate magnetic storm, the intensity of trapped radiation at high latitudes (above the peak of the outer zone) was quickly reduced about fourfold, whereas in the vicinity of the outer-zone maximum the flux of the lower-energy trapped radiation was reduced to about half its prestorm value. Narrow zones of radiation appeared at latitudes just inside the outer-zone peak, on both the day and the night side of the earth. The narrow zones merged and built up to produce a broad zone of radiation somewhat wider than the outer-zone radiation. These anomalous zones contained particles which apparently were trapped for some hours. Their intensity gradually lessened, most quickly at the lowest latitudes, over the same period that the intensity at high latitudes increased to its former value, so that within a day or two after the end of the storm the outer zone appeared normal.

The observations summarized above may be compared with the predictions of several theories [e.g., Dessler and Parker, 1959, and Piddington, 1960] that discuss the injection of particles to be trapped in the geomagnetic field. In brief, it might be considered that the narrow

zones were the trapped remnants of injected fingers of plasma, which gradually diffused owing to the longitudinal drift, so that the detailed structure disappeared as more fingers were injected, and only a broad anomalous zone remained. We believe, however, that any detailed comparison should be delayed until further events are analyzed. For this reason, also, we have made no attempt here to relate the phenomena to the zone observed by apparatus on Explorer VI [Arnoldy, Hoffman, and Winckler 1960] and on Explorer IV [Rothwell and McIlwain, 1960] between the inner and outer zones defined by Van Allen [1959].

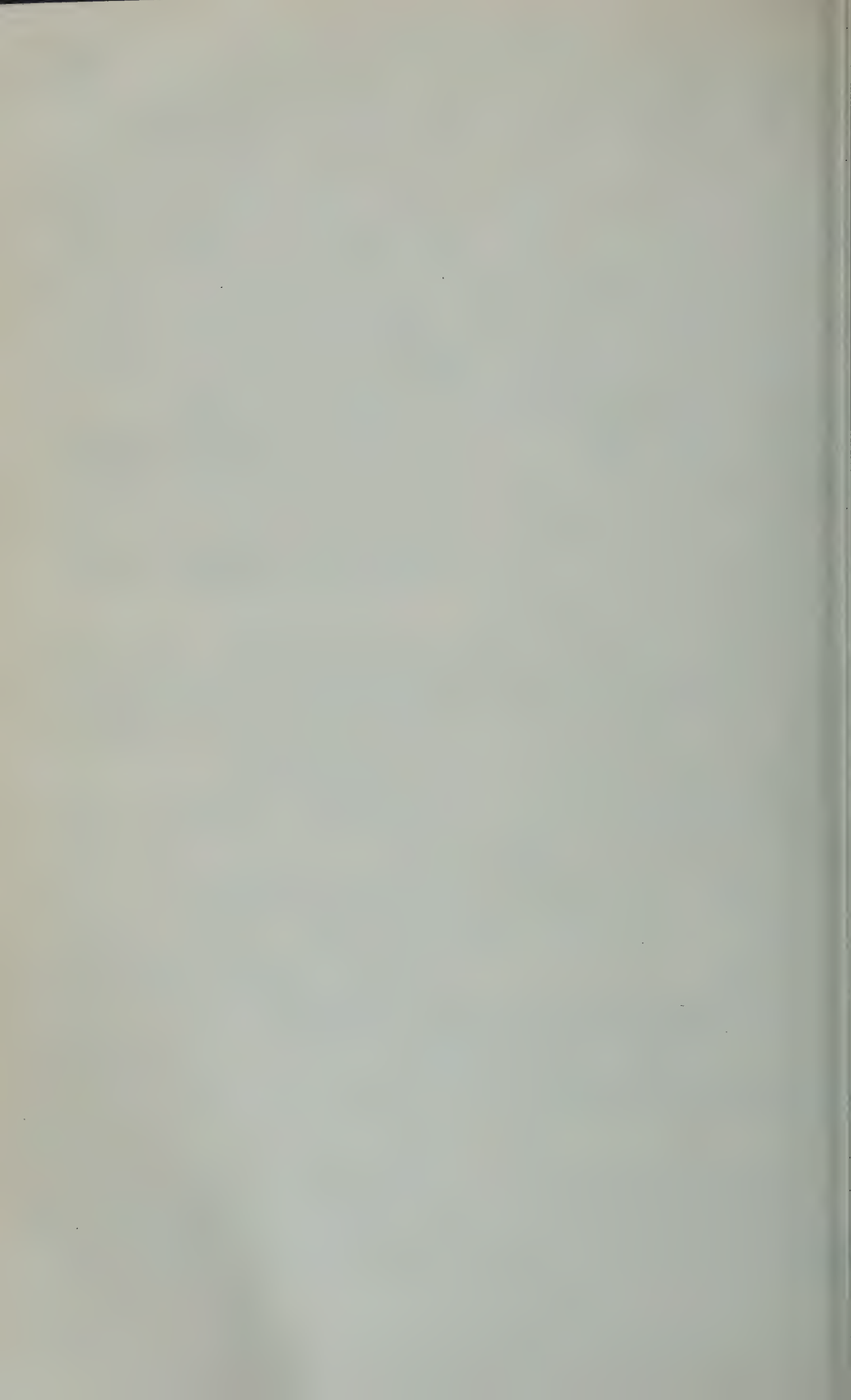
**Acknowledgment.** We are very appreciative of the useful discussions we have had with Dr. Ernest Ray of this department and with Dr. Geoff King of the National Bureau of Standards at Boulder, Colorado.

This research was assisted by the National Aeronautics and Space Administration and by the joint program of the Office of Naval Research and the Atomic Energy Commission.

#### REFERENCES

- ARNOLDY, R. L., R. A. Hoffman, and J. R. Winckler, Observations of the Van Allen radiation regions during August and September 1959, Part 1, *J. Geophys. Research*, **65**, 1361-1376, 1960.
- DESSLER, A. J., and E. N. Parker, Hydromagnetic theory of geomagnetic storms, *J. Geophys. Research*, **64**, 2239-2252, 1959.
- LUDWIG, G. H., and W. A. Whelpley, Corpuscular radiation experiment of satellite 1959 Iota (Explorer VII), *J. Geophys. Research*, **65**, 1119-1124, 1960.
- National Bureau of Standards, *Solar-Geophysical Data CRPL-F184*, Part B, 1959.
- PIDDINGTON, J. H., Geomagnetic storm theory, *J. Geophys. Research*, **65**, 93-106, 1960.
- ROTHWELL, P., and C. E. McIlwain, Scintillation counter observations from Explorer IV on trapped charged particles between the inner and outer Van Allen radiation zones, *Bull. Am. Phys. Soc.* [2], **5**, 261, 1960.
- VAN ALLEN, J. A., The geomagnetically trapped corpuscular radiation, *J. Geophys. Research*, **64**, 1683-1689, 1959.

(Manuscript received June 29, 1960.)



# Calculations of Cosmic-Ray Trajectories near the Equator

PAUL J. KELLOGG

*School of Physics, University of Minnesota  
Minneapolis 14, Minnesota*

**Abstract.** New calculations of cutoff rigidities near the equator have been carried out on a fast electronic computer, using the 48-parameter expression for the earth's magnetic field. The results are in good accord with observations.

In an earlier paper [Kellogg and Schwartz, 1958] calculations were presented of the orbits of charged particles in the earth's magnetic field and from them the position of the cosmic-ray equator was obtained. The earth's magnetic field was approximated by the sum of its dipole, quadrupole, and octupole terms. Since then a computer program has been prepared that calculates the earth's magnetic field using all the terms through  $2^{\circ}$  pole as given by Finch and others for epoch 1955. This approximation to the earth's magnetic field is much better than previously used. Figure 1 shows the 'measured' magnetic field at the geographic equator from United States charts, together with the octupole and  $2^{\circ}$  pole approximations. It will be seen that the measured field agrees with the  $2^{\circ}$  pole calculation to within 3 per cent, whereas the octupole calculation may be in error by 15 per cent. This difference is only partly due to the error of the spherical harmonic expansion; the rest is due to differences of opinion be-

tween British and United States workers about what the smoothed field should be. The calculations were carried out exactly as described in the earlier paper: the orbits of a particle arriving at the earth at a given place were calculated for various rigidities with the University of Minnesota Univac Scientific Computer. Figure 2 presents the computed position of the cosmic-ray equator together with the previously calculated position and some experimental observations. It will be seen that the new calculations are a slight improvement over the old ones. They agree with the measured position to within the combined experimental and calculational error. Figure 3 shows the calculated cutoff rigidity versus geographic longitude at the computed cosmic-ray equator.

In the first paper arguments were presented for the belief that the error in the calculated cutoff rigidity is proportional to the error in the approximation to the earth's surface field. Since

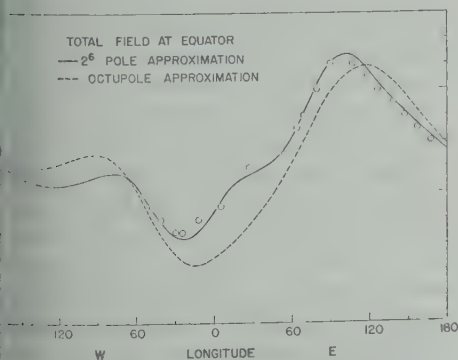


Fig. 1. The earth's magnetic field and two approximations to it. The circles are values of the field obtained from United States charts.

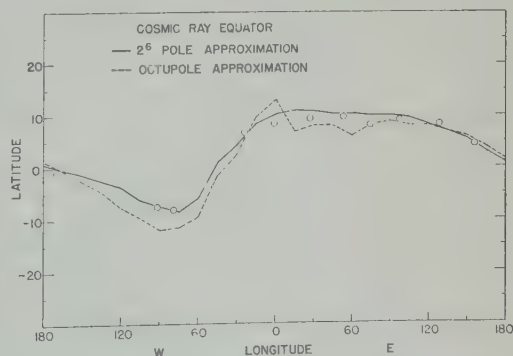


Fig. 2. The cosmic-ray equator computed using two different approximations to the earth's magnetic field. The points are measurements at airplane altitudes.



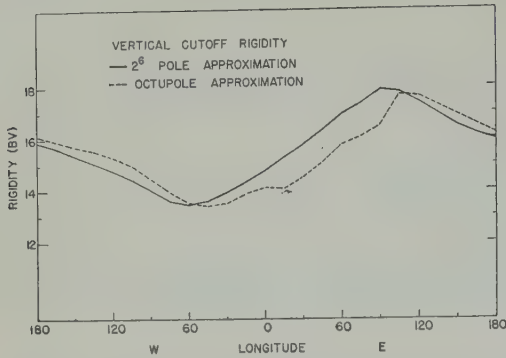


Fig. 3. Calculated cutoff rigidity at the cosmic-ray equator.

we have reduced the error in the magnetic field by a factor of 5 we should also have reduced the error in the cutoff rigidity by the same factor. Figure 3 shows that the octupole calculations were in error by as much as 1.5 bv, and so the present errors should be less than 0.3 bv.

Katz, Meyer, and Simpson [1958] have measured the cosmic-ray intensity at airplane altitudes along the cosmic-ray equator as a function of longitude. In Figure 4 we have plotted the measured cosmic-ray intensity versus calculated cutoff rigidities. All the points should fall on a smooth curve; there is some deviation, and

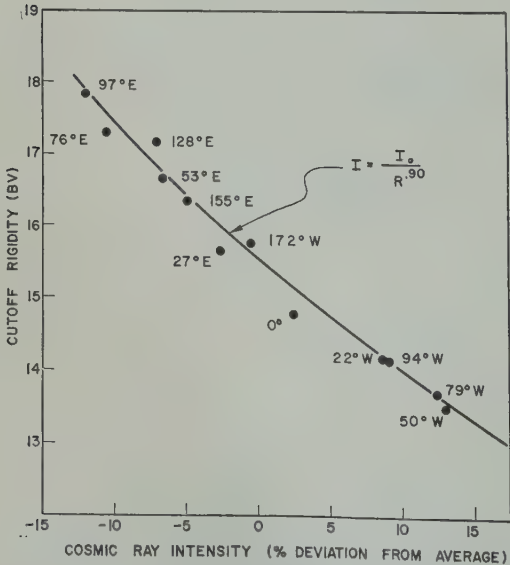


Fig. 4. Comparison of measured cosmic-ray intensity and calculated cutoff rigidity.

TABLE 1. Cutoff Rigidity in Billion Volts for Directions West of Zenith

| Zenith Angle | Longitude |       | Difference |
|--------------|-----------|-------|------------|
|              | 188°      | 27°   |            |
| 0            | 17.14     | 16.65 | 0.49       |
| 15           | 15.25     | 14.67 | 0.52       |
| 30           | 13.92     | 13.24 | 0.68       |

these seem to be slightly more than 0.3 bv. In the present paper, as in the earlier ones, the theoretical cosmic-ray equator is defined as the place along a line of longitude where the calculated rigidity in the vertical direction is a minimum. Especially at airplane altitudes, where Katz, Meyer, and Simpson's measurements were made, other parts of the sky contribute appreciably to the measured cosmic-ray intensities. In fact, cosmic-ray intensities should be a maximum at about 10° west of the zenith. Therefore, whether the contribution of the rest of the sky leads to appreciable error in the effective cutoff rigidity, calculations were carried out for directions west of the zenith for the places having the largest discrepancies, namely, the 188°E, 0°N; 27°E, 11°N; and 128°E, 53°E, 10°N.

Results for the first pair are presented in Table 1, and it can be seen that the cutoff rigidity to the west of zenith is so nearly proportional to the zenith cutoff that the mentioned discrepancies cannot be attributed to the effect of the rest of the sky. The results from the second pair are the same. It is therefore, that the calculated cutoff rigidity must actually be in error by as much as 0.3 bv. This error could be due either to failure of the 2<sup>6</sup> pole expansion to represent the earth's surface field adequately, or it could be due to the distortion of the earth's field by extraterrestrial currents. The error is only slightly more than we have estimated on the basis of error in the magnetic field, and we conclude that the magnetic field deduced from the earth's surface field gives an adequate account of cosmic-ray intensities. There is no evidence from the equatorial calculations for extraterrestrial current systems.

It would be of great interest, of course, to have some evidence in cosmic-ray measurements

TABLE 2. Cutoff Rigidity in Billion Volts at Guam ( $\theta$  = zenith angle,  $\psi$  = azimuth west of true north)

| $\theta$ | $\psi$ |       |       |                                |       |       |       |       |
|----------|--------|-------|-------|--------------------------------|-------|-------|-------|-------|
|          | 0      | 45    | 90    | 135                            | 180   | 225   | 270   | 315   |
| 0        |        |       |       | $\leftarrow 16.32 \rightarrow$ |       |       |       |       |
| 15       | 16.55  |       | 14.57 |                                | 16.60 |       | 18.77 |       |
| 30       | 17.25  |       | 13.31 |                                | 17.42 |       | 22.22 |       |
| 45       | 18.54  |       | 12.43 |                                | 18.82 |       |       |       |
| 60       | 20.52  | 14.20 | 11.86 | 13.80                          | 20.84 | 30.75 | 34.39 | 29.14 |
| 75       | 23.32  | 14.64 | 11.58 | 13.85                          | 23.83 | 39.80 | 45.49 |       |

ring current which has long been postulated as a part of magnetic storms. The Van Allen belts certainly form at least a small current. Magnetic measurements and cosmic-ray measurements of the type discussed, however, give information about the total energy content of the trapping region, and space-borne radiation counters give information only about its most energetic part.

A ring current would have a greater effect on orbits at high latitudes than on those near the equator [Ray, 1956]. Only a few orbits at high latitudes have been calculated with this method. The orbits become a great deal more complicated and have much smaller radii of curvature, so that the step size of the integration must be decreased. It takes of the order of 100 times as long to compute an orbit at high latitudes (geomagnetic latitude  $55^\circ$ ) as it does near the equator. At present we have not

enough information about the cutoff rigidity at high latitudes to be worth reporting.

Calculations of the cutoff rigidities in directions other than the zenith were made for Guam ( $13^\circ\text{N}$ ,  $145^\circ\text{E}$ ). Since these may be of use to various experimenters who took part in project Equex, they are given in Table 2.

*Acknowledgment.* This work was supported by a grant of the U. S. National Committee for the IGY.

#### REFERENCES

- Katz, L., P. Meyer, and J. A. Simpson, *Nuovo cimento, Suppl.*, 8, 277, 1958.  
 Kellogg, P. J., and M. Schwartz, Theoretical study of the cosmic ray equator, *Nuovo cimento*, 13, 761-768, 1959.  
 Ray, E. C., Effects of a ring current on cosmic radiation, *Phys. Rev.*, 101, 1142-1148, 1956.

(Manuscript received June 28, 1960.)





# Electrons of the Van Allen Radiation

PAUL J. KELLOGG

*University of Minnesota  
Minneapolis 14, Minnesota*

**Abstract.** Calculations are presented for the flux of trapped electrons to be expected on the hypothesis that these electrons are  $\beta$  decay electrons from neutrons produced in the atmosphere by cosmic rays. It is shown that such electrons can form only a very small part of the observed electron component even in the inner zone. Furthermore, the observed energy spectrum is shown to be different from that expected from neutron  $\beta$  decay, and so other sources of neutrons are not worth considering. It is concluded that the observed electrons are most probably accelerated in place by high-frequency electromagnetic fields.

In a previous publication [Kellogg, 1959a], after referred to as I, the author made estimates of the flux of particles in the Van Allen radiation belt which could be ascribed to neutrons produced in the earth's atmosphere by cosmic rays, and decaying in the storage region [Kellogg, 1958; Vernov and others, 1958]. In that paper it was tentatively concluded that protons of the Van Allen radiation could be attributed to neutron albedo, but that the electrons could not. In the meantime much more complete data have become available; it is the purpose of the present note to refine the calculations previously presented and to demonstrate conclusively that neutron albedo cannot account for the electrons. We shall concentrate on calculating the flux in the inner zone, the region whose maximum lies at about 3000 km above the earth's surface.

The calculation of the albedo contribution is done in three parts. First, we calculate the flux of neutrons out of the earth's atmosphere. From this we calculate the source density in the trap-

ping region. Then we calculate the flux of trapped particles, assuming equilibrium between injection into and loss from the storage region.

Of the neutrons that escape to decay at higher altitudes only those that have been slowed in the earth's atmosphere contribute appreciably to the trapped-electron flux. While diffusing in the atmosphere they obey the following diffusion equation:

$$\frac{\partial n}{\partial t} = \nabla \cdot (D \nabla n) + q - v \sigma_a N n + \frac{\partial}{\partial E} \left( n \frac{\partial E}{\partial t} \right) \quad (1)$$

Here  $n(r, E) d^3r dE$  is the number of neutrons having energy in  $dE$  at  $E$ , in the volume element  $d^3r$ , at  $r$ ;  $q$  is the source density, also per unit volume and energy interval. The term  $v \sigma_a N n$  is the rate of capture of neutrons by air nuclei, and the term

$$\frac{\partial}{\partial E} \left( n \frac{\partial E}{\partial t} \right)$$

represents the net flux of neutrons into the energy range  $dE$  due to collisional energy loss.  $D = \frac{1}{2} v \lambda$  is the diffusion length;  $\lambda$  is the total mean free path;  $\sigma$  and  $\sigma_a$  are the total and capture cross sections of air atoms, and  $N a$  is their number density.

If  $q$  is a function of height only, it is more convenient to write equation 1 in terms of pressure  $p$  rather than height. Assuming an isothermal atmosphere with  $p = p_0 e^{-s/h}$ ,

TABLE 1. Neutron Orbit Parameters

| Neutron Energy, ev | Time, sec | Reflection Coefficient | Maximum Height, km |
|--------------------|-----------|------------------------|--------------------|
| 0.03               | 74        | 0.86                   | 29                 |
| 0.1                | 134       | 0.76                   | 96                 |
| 0.3                | 230       | 0.62                   | 290                |
| 1                  | 420       | 0.42                   | 960                |
| 3                  | 740       | 0.22                   | 2900               |

$$\frac{\partial}{\partial z} \left( D \frac{\partial n}{\partial z} \right) = \frac{1}{3} \frac{v \lambda p^2}{h^2} \frac{\partial^2 n}{\partial p^2} \quad (2)$$

where we have used the fact that  $\lambda p$  is independent of pressure.

To give an explicit solution of equation 1 we need the boundary condition on  $n$  at the top of the atmosphere. In I, the approximation was made that neutrons leaving the atmosphere never returned. In the present calculation we wish to take the effect of gravity into account. If a neutron has insufficient energy to escape the earth's gravitational field there is a chance that it will return to the atmosphere before decaying. Thus the top of the atmosphere has a certain reflectivity  $R$ , which depends on the neutron energy. Values of  $R$  computed on the approximation of a plane earth are given in Table 1 together with the time of flight above the atmosphere and the maximum height reached. The results are not much different from the actual results so long as the neutron energy is below about half of the escape energy of 0.7 ev.

In terms of this reflection coefficient, the boundary condition at the top of the atmosphere becomes

$$n + \frac{2}{3} \lambda \left( \frac{1+R}{1-R} \right) \frac{\partial n}{\partial z} \Big|_{p=0} = 0 \quad (3)$$

As usual in diffusion problems, we satisfy this boundary condition approximately by requiring  $n$  to vanish at a pressure  $p_1$  given by

$$p_1 = -\frac{2}{3} \left( \frac{1+R}{1-R} \right) \frac{\lambda p}{h} \quad (4)$$

Then the solution of equation 1 for equilibrium is

$$n = \int dp_0 q(p_0) \left( \frac{\lambda p}{E v} \right) \frac{M}{2 \sqrt{\pi m k L}} \cdot \left\{ \exp \left[ -\frac{(p-p_0)^2}{L^2} \right] - \exp \left[ -\frac{(p+p_0-2p_1)^2}{L^2} \right] \right\} \cdot \exp \left[ -\sqrt{\frac{\alpha}{E}} \right] \quad (5)$$

where

$$L^2 = \frac{2M}{3m} \int_E^{E_0} \left( \frac{\lambda p}{h} \right)^2 \frac{dE}{E} = L^2(E)$$

is the diffusion length, and

$$\alpha = \left( \frac{\sigma_a v}{\sigma} \right)^2 \frac{M^2}{2m} \simeq 0.05 \text{ ev}$$

and  $q dp_0$  is the number of neutrons produced per square centimeter per second in the pressure interval  $dp_0$ . We have tacitly assumed that neutrons were produced at energy  $E_0$ .  $E_0$  should represent the average energy of evaporated neutrons in stars, approximately 2 Mev. But this region  $\lambda$  is quite large, and therefore is quite sensitive to  $E_0$ , although intuitively it is apparent that the choice of  $E_0$  should not be important. We shall be able to present the result in a form relatively independent of  $E_0$ .

The upward flux of neutrons of energy  $E$  from the top of the atmosphere is, according to diffusion theory,

$$J(E) = \frac{1}{3} \frac{v \lambda}{1-R} \frac{\partial n}{\partial z} \Big|_{p=0} = \frac{1}{1-R} \frac{v \lambda p}{3 h} \frac{\partial n}{\partial p} \Big|_{p=0}$$

We now proceed to calculate the rate at which these neutrons decay in the trapping region. We assume that all the neutrons described in equation 8 are traveling straight upward. Thus the actual nearly isotropic distribution is taken into account only through the factor  $1/3$ .

The number that decay at a height  $z$  in range  $dz$  is

$$Q(E) = J(E) P(E, z) = J(E) e^{-t/\tau} (1/v\tau) dz$$

where  $t$  is the time the neutron takes to reach height  $z$ ,  $\tau$  is the neutron mean life,  $v$  is the vertical component of velocity at height  $z$ ,  $P$  is the probability that a neutron of energy  $E$  decays in  $dz$ .  $t$  and  $z$  are given by

$$z = v_0 t - \frac{1}{2} g t^2 \quad v = v_0 - g t$$

where  $v_0$  is the initial velocity and  $g$  the acceleration of gravity. For neutrons that can escape there are two solutions for  $t$ , corresponding to

to neutrons traveling upward, and those falling back; and the right-hand side of (9) should be the sum of terms for each solution. For large  $Pv_0\tau$  approaches 1. The total number of decays per unit volume and time is therefore given by

$$= \int_{mg}^{\infty} Q(E) dE$$

$$= \int_{mg}^{\infty} \frac{1}{1-R} \left( \frac{v_0 \lambda p}{3h} \right) P \frac{\partial n}{\partial p} \bigg|_{p=0} dE \quad (11)$$

We still must evaluate the atmospheric source density  $q(p_0)$  appearing in equation 5. We do this as in I, by using the results of *Soberman* [1966], who measured the counting rate in  $\text{BF}_3$  proportional counters at balloon altitudes, that measured

$$C = N_B \int_0^{\infty} \sigma_B v n dE \quad (12)$$

where  $N_B$  is the total number of  $\text{B}^{10}$  nuclei in the counters, and  $\sigma_B$  is the  $\text{B}^{10}$  capture cross section.  $\sigma_B v$  is independent of energy, so we can write

$$\frac{dC}{dp} \bigg|_{p=0} = N_B \sigma_B v \int_0^{\infty} \frac{\partial n}{\partial p} \bigg|_{p=0} dE \quad (13)$$

The integrals of equations 11 and 12 differ only in that the integral of (11) has (a) a nonzero

lower limit and (b) an additional factor

$$(P/(1-R)) (v_0 \lambda p / 3h) \quad (14)$$

under the integral sign. This factor approaches a constant  $\lambda p / 3h\tau$  at large velocity. In Figure 1 it is plotted (normalized to unity) for two different heights, 1000 and 3000 km.

There is therefore a very close relation between the measured quantity  $dC/dp$  and the decay rate  $Q$ . The diffusion theory which has been presented serves only to establish this relation, and our results are independent of most of the parameters. The result required from diffusion theory is the energy spectrum of the diffusing neutrons, which on our simple assumptions is roughly proportional to

$$e^{-\sqrt{\alpha/E}} E^{-3/2} \quad (15)$$

This function has a maximum at

$$E = \alpha/9 \simeq 0.005 \text{ e.v.}$$

This cannot be quite correct, because the maximum in the energy spectrum cannot fall below thermal energy and this is nearly 10 times higher. *Poole, Nelkin, and Stone* [1958] have made more careful calculations of the energy spectrum using an integral expression for the collision energy loss, but insertion of their energy spectrum into equations 11 and 13 does not make any great difference in the results.

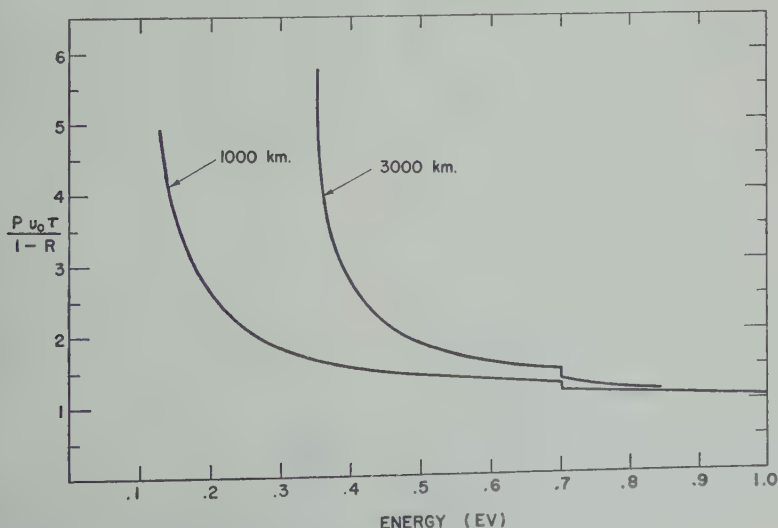


Fig. 1. Relative probability of neutron decay as a function of energy, for two altitudes.



We divide equation 11 by equation 13 to give

$$\frac{N}{(dc/dp)} = \frac{\lambda p}{3h\tau N_B \sigma_B v} \frac{1}{x} \quad (16)$$

where

$$x = \frac{\int_{m\phi}^{\infty} \frac{Pv_0\tau}{1-R} \frac{\partial n}{\partial p} dE}{\int_0^{\infty} \frac{\partial n}{\partial p} dE} \quad (17)$$

would be unity if there were no gravity and negligible neutron decay before height  $z$ , and is not greatly different from unity. For a height of 1000 km,  $x = 1.1$ , and for 3000 km,  $x = 0.6$ .

The evaluation of  $dc/dp$  at the top of the atmosphere involves some extrapolation from the highest measured values. The neutron density  $n$  is to be extrapolated to zero at a pressure  $p_1$  given by equation 4,  $p_1$  depends on energy, so it is not clear at what pressure  $c$  should vanish. If we take the pressure  $p_1$  corresponding to the maximum of the energy spectrum  $E = 0.005$  ev,

$$\begin{aligned} p_1(E = 0.005) \\ &= -[(1+R)/(1-R)]2.2 \text{ mb} \\ &= -23 \text{ mb} \end{aligned}$$

the factor  $(1+R)/(1-R)$  is quite large. The more careful calculations by Poole, Nelkin, and Stone indicate that the maximum in the energy spectrum comes at about 0.05 ev, where the factor  $(1+R)/(1-R)$  is only about 2.5 and  $p_1 = -5$  mb. This extrapolation agrees better with the slope of Soberman's measured curves. Since Soberman's highest measurements were made at about 20 mb, we will not be in error by more than 25 per cent if we extrapolate to -10 mb. Then, at the top of the atmosphere near the equator, Soberman's measurements give  $dc/dp = 0.25$  count/min-mb. The decay rate is then  $N = 1.1 \times 10^{-12} x$  decay/cm<sup>2</sup> sec over the equator, where  $x$  is given in equation 17. The calculation we have presented is presumably similar to the one referred to by Hess [1959] as a multigroup diffusion calculation. The neutron production rate  $q$ , which we have evaluated using Soberman's measurements, was evaluated by Hess from his aircraft measurements.

*Equilibrium flux in inner zone.* A significant fraction of the electrons whose decay rate we have calculated will be emitted at pitch angles great enough so that they will be held in the denser parts of the earth's atmosphere. In the converging lines of the earth's magnetic field. There they will be trapped until they have lost their energy by collisions with other electrons. But although ionization loss is the dominant loss mechanism (assuming that the magnetic field is stable and regular enough so that collisions are the main cause of loss), Christy [1959] has pointed out that ionization loss is greatly accelerated by the scattering of electrons into orbits which penetrate into the denser parts of the atmosphere. Then they make more collisions near the lowest parts of their orbits and in these can never be scattered into orbits with higher turning points, but only into still lower ones. Thus once their orbits penetrate into the denser part of the atmosphere they are eventually lost.

We therefore neglect energy loss in the trapping region and compute only the likelihood against scattering into the denser part of the atmosphere.

Actually, neglect of energy loss is not a good approximation in the region we are considering. The mean-square scattering angle [Rossi, section 216] is approximately, for electrons in hydrogen,

$$\theta_s^2 = 12(mc/\beta p)^2 \text{ (radian)}^2 \text{ cm}^2/\text{g}$$

The energy loss is about  $(4/\beta^2)$  Mev/g/cm distance traversed while an electron of energy  $E$  loses half its energy is then

$$E\beta^2/(16mc^2)$$

and in this distance

$$\theta_s^2 = \frac{3}{4}mE/p^2 \text{ (radian)}^2$$

For energies up to  $mc^2$ , such as we are considering, scattering and energy loss are of roughly equal importance. By neglecting energy loss we obtain only an upper limit to the lifetime, and therefore an upper limit to the flux of trapped electrons.

To calculate the number of electrons in the inner Van Allen zone, according to this picture, we use the Fokker-Planck equation

density of particles as a function of position angle along velocity and magnetic field.  $\psi, r) \sin \psi d\psi dV$  be the number of electrons of energy  $E$  in volume  $dV$  and pitch angle  $\psi$ . We use the now well-known perturbation theory of particles trapped in a magnetic mirror (Schep, 1929; *Alfvén*, 1950] and neglect diffusion across lines of field. In a time  $dt$ ,  $f$  has the following changes: (1) The converging lines of field give an effective force along the field which moves particles in angle elements adjacent to  $d\psi$ , and vice versa. (2) Particles in volume elements adjacent to  $dV$  move into  $dV$ , and particles in  $dV$  move out. (3) Particles in  $dV$  and  $d\psi$  may be scattered into different pitch angles, and vice versa. (4) Neutrons decay, emitting  $\beta$  particles into  $dV$  and  $d\psi$ . The first term may be evaluated in the usual way for the Boltzmann equation using the fact that the force along the field is given by

$$F = -\frac{mv^2 \sin^2 \psi}{2B} \frac{dB}{dl} \quad (20)$$

The second term may be evaluated similarly using the fact that the velocity along the line of field is  $v \cos \psi$ . The third term has been evaluated by *Rosenbluth, MacDonald, and Judd* (1958) to be, for the present case,

$$\frac{1}{T} \frac{\partial}{\partial \mu} \left[ (1 - \mu^2) \frac{\partial f}{\partial \mu} \right] \quad (21)$$

where  $\mu = \cos \psi$ . In this we have assumed that scattering is due to particles so heavy that the electron loses no energy. This means that we neglect scattering by electrons which have pitch angles smaller by a factor of 4 because of their reduced mass. The effective lifetime  $T$  is given by

$$\frac{1}{T} = \frac{2\pi e^4 N_a (1 - \beta^2)}{m^2 v^3} \ln \left( \frac{2}{\theta_m} \right) \quad (22)$$

where  $\theta_m$  is the distance of closest approach/Debye

$$= \frac{mv^2}{2e^2 [kt/(4\pi N_a e^2)]^{1/2}}; \quad \ln \left( \frac{2}{\theta_m} \right) \approx 28$$

where  $\theta_m$  has been evaluated for  $kt = 0.1$  ev,  $10^{14}$  atoms/cm<sup>3</sup>,  $1/2 mv^2 = 330$  kev, and

the logarithm is not sensitive to reasonable changes of these quantities.

Equation 22 has been modified to take into account relativistic effects. The full Fokker-Planck equation is then

$$\begin{aligned} \frac{\partial f}{\partial t} = & -\frac{v}{2B} \frac{dB}{dl} \frac{\partial}{\partial \psi} (f \sin^2 \psi) \\ & - v \sin \psi \cos \psi B \frac{dB}{dl} \frac{\partial}{\partial B} \left( \frac{f}{B} \right) \\ & + \frac{1}{T} \frac{1}{\sin \psi} \frac{\partial}{\partial \psi} \left( \sin \psi \frac{\partial f}{\partial \psi} \right) + \frac{N}{4\pi} \rho(E) \end{aligned} \quad (23)$$

where the terms on the right are written in the order in which they were described.  $\rho(E)$  is a factor proportional to the neutron  $\beta$  decay spectrum and normalized so that

$$\int \rho dE = 1$$

The first two terms of equation 23 describe the trapping; the last two describe injection into and scattering out of the trapping region. Since an electron is reflected many times by the magnetic field before it is scattered out, the first two terms will be very much larger than either of the second two, and so we consider the equation

$$\begin{aligned} \frac{\partial f}{\partial t} = & -\frac{v}{2B} \frac{dB}{dl} \frac{\partial}{\partial \psi} (f \sin^2 \psi) \\ & - v \sin \psi \cos \psi B \frac{dB}{dl} \frac{\partial}{\partial B} \left( \frac{f}{B} \right) \end{aligned} \quad (24)$$

A general solution to this equation is given by

$$f = G(\sin^2 \psi/B)$$

where  $G$  is an arbitrary function of its argument. Now we attempt to solve the full equation 23 by specifying this arbitrary function. This will be possible if we make a simple enough approximation to atmosphere density. It is generally assumed that the earth's atmosphere is mainly nitrogen and oxygen to a height of about 1000 to 1500 km, and hydrogen above that. There is a corresponding increase in scale height, so that, above a height which we will call 1000 km, atmospheric density changes only rather slightly, and *Johnson* [1960] estimates that, between the lower boundary of the Van Allen zone

at about 1000 km and its maximum at about 3000 km, the density of the atmosphere changes by only a factor of 2. We may reasonably take atmospheric density as constant and equal to  $10^4$  H atoms/cm<sup>3</sup>. Below 1000 km the density increases sharply, and by our previous argument electrons whose turning points lie in this region are rapidly lost. For this reason we take the atmospheric density below 1000 km to be infinite. This implies that the function  $G$  vanishes if  $\sin^2 \psi/B$  corresponds to particles which have turning points below the 1000-km level. A solution of equation 23 which satisfies this requirement is

$$G = (1/8\pi)NT\rho \ln [(B_1/B) \sin^2 \psi] \quad (25)$$

where  $B_1$  is the value of  $B$  at the 1000-km level. The total flux of electrons of all energies is given by

$$\int \rho T dE \frac{N}{8\pi} \ln \left( \frac{B_1}{B} \sin^2 \psi \right) \quad (26)$$

$\int \rho T dE$  is the average lifetime of  $\beta$  decay electrons against scattering through some angle of the order of 1 radian. If the spectrum of neutron  $\beta$  decay is the allowed spectrum

$$\rho(E) \propto \sqrt{E^2 - m^2 c^4} (E_{\max} - E)^2 E \quad (27)$$

where  $E$  is now total energy, then the average lifetime  $\int \rho T dE$  in an atmosphere of  $10^4$  H atoms/cm<sup>3</sup> is  $3.3 \times 10^8$  sec.

The logarithm has its maximum value at the equator when the flux is measured perpendicular to  $B$ , and is then  $\ln (B_1/B_0)$ , where  $B_0$  is the value of  $B$  at the equator and  $B_1$  is the value of  $B$  at the 1000-km level on the same line. For the line passing 3000 km above the equator,

$$\ln (B_1/B_0) \approx 0.96$$

so the maximum flux of electrons attributable to neutron  $\beta$  decay is, at 3000 km,

$$\begin{aligned} & \frac{1}{8\pi} N \int \rho T dE \left( x \ln \frac{B_1}{B_0} \right) \\ &= 3.5 \times 10^5 \left( x \ln \frac{B_1}{B_0} \right) \\ &= 2.5 \times 10^5 \text{ electrons/sterad cm}^2 \text{ sec} \end{aligned}$$

The average lifetime for storage in the heart

of the inner zone is given by

$$\ln \frac{B_1}{B_0} \int \rho T dE = 3.3 \times 10^8 \text{ sec}$$

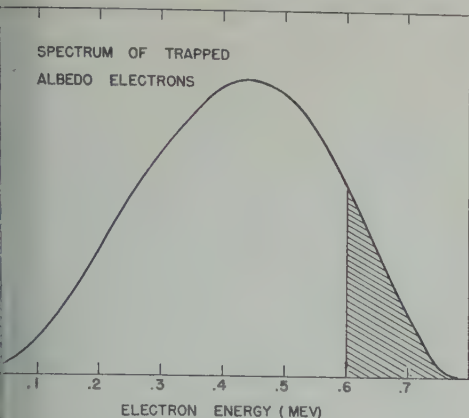
and corresponds to that of a 325-kev electron. In this time, such an electron will travel  $3.3 \times 10^8$  g/cm<sup>2</sup> in the atmosphere we are considering. The range of such an electron, however, is about 0.07 g/cm<sup>2</sup>, and so scattering and energy loss are of roughly equal importance, and the scattering lifetime is about twice too long.

According to Van Allen [1959], the maximum flux observed at this height is actually 2 electrons/sterad cm<sup>2</sup> sec, to within a factor of 2. It is difficult to see how the approximations that have been made in this calculation could introduce an error as great as a factor of 2. There is also, of course, some uncertainty in the density of the exosphere. Nevertheless, even a no reasonable revision of the present estimates will allow agreement of the measured and calculated fluxes. This is even more strongly true for the outer zone than for the inner one. In factors that go into the flux, only  $\ln(B_1/B_0)$  will be larger, and that not much larger, for the outer zone. The measured flux is much greater, however, and large fluctuations in intensity make it clear that the lifetime is much smaller than we have calculated. We conclude that cosmic-ray-produced neutron albedo contributes only a very small part of the electron flux in the Van Allen radiation belts.

We may expect additional contributions to neutron decay from two sources: high-altitude nuclear explosions, and large fluxes of high-energy protons from the sun. The lower-energy proton fluxes are confined to the polar cap, and some of the neutrons produced may escape and decay in the region of the inner zone. It is difficult to estimate this contribution, but it will not be likely to be as much as the contribution of the cosmic-ray effect. High-altitude nuclear explosions produce many neutrons and are capable of making a very large contribution to the trapped radiation. However, the earliest high-altitude explosion known to the author occurred on August 1, 1958, and the electron flux in the inner zone a few days before that was at a minimum as ever.

Strong arguments against neutron  $\beta$





2. Spectrum of trapped electrons which would be produced by neutron  $\beta$  decay.

An important source of the radiation belt provided by measurements of the electron spectrum. We discuss two such measurements, first by *Van Allen, McIlwain, and Ludwig [9a]* with Explorer IV, and the second by *Holly and Johnson [1960]*. The expected spectrum of trapped electrons from neutron decay is shown in Figure 2. It is peaked at about 450 keV and it drops rapidly at higher energies because the  $\beta$  decay spectrum does and at lower energies because the lifetime of an electron is proportional to  $v^3$ . The satellite Explorer IV measured, among other things, a scintillator which measured the total energy flux of electrons of energy greater than 20 keV, and another which counted electrons above 600-keV energy. For example, measurements at altitudes about 2000 km over South America where the inner radiation belt dips lowest gave  $76 \text{ erg/cm}^2 \text{ sec sterad}$  for the energy flux and  $< 10^8 \text{ electrons/cm}^2 \text{ sec sterad}$  for the flux of electrons having energies above 600 keV. If the electrons had the spectrum shown in Figure 2, a flux of  $76 \text{ ergs/cm}^2 \text{ sec sterad}$  would correspond to a total flux of  $1.1 \times 10^8 \text{ electrons/cm}^2 \text{ sec}$ , of which  $1.3 \times 10^7$  would have energy above 600 keV. The observed counting rate is about 5 times lower, so that the actual spectrum is flatter in low-energy electrons. The observed discrepancy between the two detectors we have discussed varies by a factor of 2 either way from the value we have given, but this is not enough to account for the discrepancy, and, further-

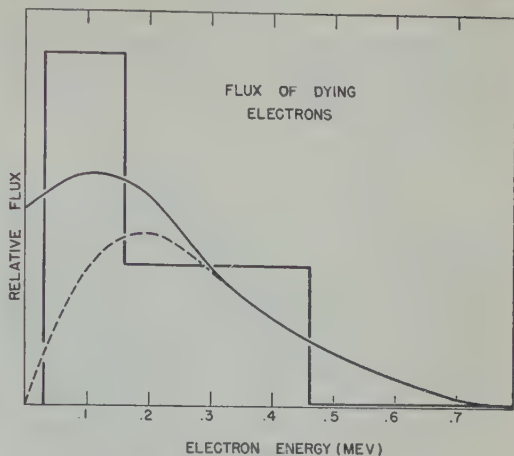


Fig. 3. Comparison of measured flux of electrons out of the trapping region with that expected on the neutron albedo hypothesis.

more, the variation itself is an argument against neutron  $\beta$  decay as the main source.

Holly and Johnson have made measurements of the electron spectrum which are more detailed but more difficult to interpret. They measured the spectrum of the trapped electrons at an altitude of about 780 km. This height is not sufficient to be outside the nitrogen-oxygen atmosphere, and this leads to some difficulty of interpretation. Since Holly and Johnson's measurements were not made in the trapping region, the electrons that they see are dying. If we neglect the energy loss during their lifetime in the trapping region, the flux of dying electrons out of the trapping region should be the same as the  $\beta$  decay spectrum and is not complicated by any factor for storage lifetime.

However, the electrons are dying by ionization loss, which distorts the  $\beta$  decay spectrum slightly. We may take the energy loss of an electron per unit time to be inversely proportional to  $v$ , which corresponds to taking the energy loss per  $\text{g/cm}^2$  to be inversely proportional to  $v^3$ . Then  $n dE$ , the relative number of electrons in the energy interval  $dE$ , obeys the equation

$$\frac{\partial n}{\partial t} = \frac{\partial}{\partial E} \left( n \frac{k}{v} \right) + \rho(E) \quad (28)$$

where  $k/v = dE/dt$  and  $\rho$  is the  $\beta$  decay spec-

trum as before. The first term on the right represents the net increase in number in  $dE$  due to energy loss, and the second term injection by  $\beta$  decay. The solution of this equation is, for  $\partial n/\partial t = 0$ ,

$$n = \frac{v}{k} \int_E^{E_{\max}} \rho(E') dE' \quad (29)$$

This spectrum, plotted as a dotted curve in Figure 3, refers to measurements deep in the atmosphere. As the altitude is increased, the spectrum should change continuously from that of Figure 3 to one corresponding to Figure 2, modified by a factor of  $v$  to give flux instead of number. But Figure 3 represents the spectrum which is pushed toward low energies by the maximum possible amount. All higher-altitude spectra must have relatively more high-energy electrons. The correct spectrum as a function of altitude could only be obtained by a solution of the Fokker-Planck equation with terms for energy loss and a better model of the atmosphere; but the limiting spectrum represented by Figure 3 will suffice to rule out neutron  $\beta$  decay electrons as a significant source.

While circulating in the trapping region, electrons and also protons may make close collisions with electrons and give them considerable energy. The struck electrons will then also form part of the measured flux. Electrons, however, are rather inefficient producers of such  $\delta$  rays. For a typical electron of 300-kev energy, about 1/40 of the total energy loss goes into  $\delta$  rays of energy greater than 20 kev [Rossi, 1952], and these rays have a spectrum proportional to  $1/E^2$ . The protons are somewhat more efficient at producing  $\delta$  rays, but their flux is very much smaller.

The total spectrum of dying  $\beta$  decay electrons and electron-produced  $\delta$  rays is given by the solid curve in Figure 2, together with a histogram of Holly and Johnson's measurements normalized to give the same flux in the interval 160 to 460 kev. The calculations agree with experiment moderately well in the lower energy interval, but about 10 times too many electrons are predicted for the upper energy interval. It is clear that the measured spectrum falls much more steeply with increasing energy than the predicted one, again indicating that another source for these electrons must be found.

At the present time we know enough about the Van Allen zone to place stringent requirements on a source of electrons in the zone. We know from the stability and stancy of the Argus shell [Van Allen, McIlwain, and Ludwig, 1959b] that at least during magnetically quiet times electrons do not move from one magnetic shell to another. Further, it appears that Argus electrons were detected only after a magnetic storm and that, therefore, magnetic storms cause only slight diffusion of electrons from one shell to another although they do affect the intensity. It is nearly certain that the protons in the inner zone are not as albedo neutrons [Hess, 1959; Freden, White, 1959 and 1960], that their lifetimes are governed by ionization losses and nuclear collisions, and that their intensity is relatively unaffected by magnetic storms. Electrons created in the atmosphere by various processes such as photoelectric ejection by solar X-rays would have very short lifetimes since their turning points would almost always lie in the dense atmosphere. Therefore, it appears that the protons in the inner zone must be accelerated in place by some mechanism which must not disturb the protons.

The author [Kellogg, 1959b] has discussed the effect of low-frequency electromagnetic fields on trapped particles. In a static field varying slowly in space the orbit of a particle has constants of the motion: the energy  $W$ , the magnetic moment  $(W \sin^2 \psi)/B$ ; and the momentum analogue  $\int v_{\parallel} dl$ . Time-varying fields whose period is long compared with the cyclotron frequency and the time the particle takes to go between turning points will change  $W$  only. A particle whose  $W$  changes alone will move in toward or away from the earth contrary to observation. Therefore the period of the accelerating fields must be shorter than the time between turning points.

In the same paper the author showed that low-frequency fields accelerate particles so that their fractional energy change is constant; therefore low-frequency fields would have a larger effect on the protons, which have energies around 50 Mev, than on the electrons which have energies of only 50 kev.

The argument is as follows: The rate of energy gain of a particle is given by  $eE \cdot v$ ,

electric field is slowly varying, gyrations change to zero and so  $v$  is the drift velocity. Low-frequency electric fields will be expected to be perpendicular to the earth's magnetic field because the conductivity parallel to the magnetic field is so great. Therefore only the perpendicular component of drift velocity enters, and this has two terms proportional respectively to  $\sin^2 \psi$  and  $W \cos^2 \psi$ , where  $W$  is the energy  $\psi$  is the pitch angle [Gunn, 1929; Alfvén, 1959]. For this argument the electric field must be perpendicular to the magnetic field and of frequency small compared with the cyclotron frequency. There is another argument that important electric fields ought to be perpendicular to the earth's magnetic field. Electric fields parallel to the magnetic field are inefficient for acceleration because particles have lower turning points when they are accelerated and so are lost before gaining much energy. If the electric field is perpendicular to the magnetic field, then, even for periods comparable to the time between turning points (of order of seconds), the rate of energy change is proportional to the energy. In order to accelerate electrons without disturbing the protons, then, even higher-frequency fields in the trapping region are required. This conclusion, however, is not as firm as the first. Frequencies above the electron cyclotron frequency would be unfavorable. It is predicted that there are instabilities in plasmas which produce fields of order the electron cyclotron frequency [Rosenbluth, 1959], so these fields could be generated by collective effects of electrons of lower energy than those observed by the radiation counters. The spectrum of electrons in the inner zone is not too dissimilar from the spectrum in the outer zone or zones. Further, recent observations in the space probe, Pioneer V (J. R. Jokipii, in conversation), make it appear that electrons in the outer zone are also locally generated, as they do not appear in great numbers in the space probe at a time when the density in the outer zone increases greatly. The observations point to a common origin for the electrons in the inner and outer zones.

## REFERENCES

- Alfvén, H. *Cosmical Electrodynamics*, Oxford University Press, New York, 237 pp., 1950.
- Christofilos, N. C., The Argus experiment, *J. Geophys. Research*, **64**, 869-876, 1959.
- Freden, S. C., and R. S. White, Protons in the earth's magnetic field, *Phys. Rev. Letters*, **3**, 9-11, 1959.
- Freden, S. C., and R. S. White, Particle fluxes in the inner radiation belt, *J. Geophys. Research*, **65**, 1377-1383, 1960.
- Gunn, R., An electromagnetic effect of importance in solar and terrestrial magnetism, *Phys. Rev.*, **33**, 832, 1929.
- Hess, W. N., Van Allen belt protons from cosmic ray neutron leakage, *Phys. Rev. Letters*, **3**, 11-13, 1959.
- Holly, F. E., and R. G. Johnson, Measurements of radiation in the lower Van Allen belt, *J. Geophys. Research*, **65**, 771-772, 1960.
- Johnson, F. S., The ion distribution above the  $F_2$  maximum, *J. Geophys. Research*, **65**, 577-584, 1960.
- Kellogg, P. J., Possible explanation of the radiation observed by Van Allen at high altitudes in satellites, *Nuovo cimento*, **11**, 48-66, 1959a.
- Kellogg, P. J., Van Allen radiation of solar origin, *Nature*, **183**, 1295-1297, 1959b.
- Poole, Nelkin, and Stone, *Progr. in Nuclear Energy*, **11**, 91, 1958.
- Rosenbluth, M. N., Recent theoretical developments in plasma stability, *Bull. Am. Phys. Soc.*, *ser. II*, **4**, 197, 1959.
- Rosenbluth, M. N., W. M. MacDonald, and D. L. Judd, Fokker-Planck equation for an inverse-square force, *Phys. Rev.*, **107**, 1-6, 1957.
- Rossi, B. *High Energy Particles*, Prentice-Hall, New York, 569 pp., 1952.
- Singer, S. F., 'Radiation belt' and trapped cosmic-ray albedo, *Phys. Rev. Letters*, **1**, 171-173, 1958.
- Soberman, R., High altitude cosmic ray neutron intensity variations, *Phys. Rev.*, **102**, 1399, 1956.
- Van Allen, J. A., The geomagnetically trapped corpuscular radiation, *J. Geophys. Research*, **64**, 1683-1689, 1959.
- Van Allen, J. A., C. E. McIlwain, and G. H. Ludwig, Radiation observations with satellite 1958e, *J. Geophys. Research*, **64**, 271-286, 1959a.
- Van Allen, J. A., C. E. McIlwain, and G. H. Ludwig, Satellite observations of electrons artificially injected into the geomagnetic field, *J. Geophys. Research*, **64**, 877-892, 1959b.
- Vernov and others, *Rocket and Satellite Symposium*, Fifth CSAGI Meeting, Moscow, 1958.

(Manuscript received May 26, 1960.)





## Geomagnetic Storm Sudden-Commencement Rise Times

A. J. DESSLER, W. E. FRANCIS, AND E. N. PARKER<sup>1</sup>

*Lockheed Aircraft Corporation, Missiles and Space Division  
Palo Alto, California*

**Abstract.** Adopting the view that the sudden commencement of a geomagnetic storm (SC) is the result of the impact on the geomagnetic field of an abrupt solar-plasma front, the form of the SC observed on the surface of the earth is investigated. A model is constructed to represent the shape of the geomagnetic field boundary as perturbed by the solar plasma. Calculations (carried out in the equatorial plane for simplicity) show that, regardless of how abrupt may be the impact of a solar-plasma front on the geomagnetic field, the variation in hydromagnetic transit times from different positions on the boundary, down to a point on the surface of the earth, yields SC rise times of several minutes at ground level. These times are in agreement with the observed SC rise times.

**Introduction.** A geomagnetic storm typically starts with an initial impulse in  $H$ , the horizontal component of the earth's magnetic field. This impulse in  $H$  is called the sudden commencement (SC). In low and temperate latitudes, the change in  $H$  is almost always positive. At other latitudes, the initial impulse of the SC is often negative. The SC amplitude is typically about  $30 \gamma$  ( $1 \gamma = 10^{-5}$  gauss).

The mechanism whereby plasma coming from the sun could perturb the geomagnetic field so as to account for some of the features of a SC was developed by *Chapman and Ferraro* [1931, 1932]. Since at that time interplanetary space in the region above the ionosphere were thought to be a vacuum, the hydromagnetic aspects of the SC problem were not considered. In this paper, the work of *Chapman and Ferraro* is extended to include hydromagnetic effects.

The possible importance of hydromagnetic effects in propagating magnetic disturbances to the earth's surface was first pointed out by *Green* [1957]. The propagation of SC's by hydromagnetic waves has been investigated quantitatively on the basis of an idealized model by *Francis, Green and Dessler*, 1959].

One of the major features of the SC which has not yet been satisfactorily explained is the

rise time of the initial change in  $H$  (i.e., the SC rise time). The SC rise time is observed to lie generally within a range of 1 to 6 minutes. These times are long compared with the rise time of the solar-plasma front as it moves through interplanetary space or with the time required for the plasma front to sweep by the earth. The 1- or 2-day delay between a solar flare and a terrestrial SC indicates that the velocity of the solar plasma is of the order of  $10^8$  km/sec. Theoretical considerations of the rise time of the plasma front lead to the conclusion that it is probably not greater than 10 seconds [*Dessler and Parker*, 1959]. The time required for the plasma front to sweep by the geomagnetic field is only about 30 seconds. The problem, then, is to establish the mechanism whereby this relatively sharp impact of the solar plasma on the geomagnetic field is stretched out in time by about one order of magnitude before it is observed on the earth's surface.

A previous suggestion that the long SC rise times might be accounted for by the attenuation of the higher-frequency Fourier components of the hydromagnetic waves which produce the SC [*Dessler and Parker*, 1959] has been shown to be quantitatively incorrect; waves with periods longer than about 10 seconds will not be appreciably attenuated, as shown by calculations of the attenuation of hydromagnetic waves in the ionosphere [*Francis, Dessler, and Karplus*, 1960]. Thus, hydromagnetic waves with rise

<sup>1</sup>Permanent address: Enrico Fermi Institute for Nuclear Studies, University of Chicago, Chicago, Illinois.

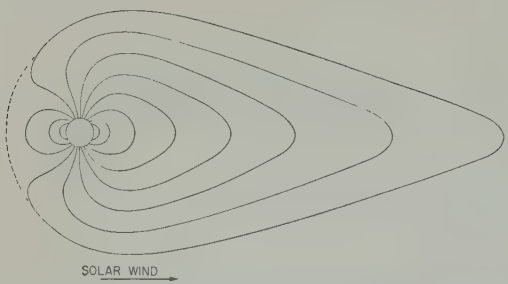


Fig. 1. Sketch showing the probable distortion of the geomagnetic field by the solar wind (after Johnson, 1960a).

times greater than 10 seconds will not be significantly altered on passing through the lower ionosphere.

In the following section, a simple hydromagnetic model is presented which is capable of accounting for the observed SC rise times. It will be assumed that the plasma-front rise time is negligibly small. Only positive SC's are considered. The negative SC often observed at high-latitude stations is not explained.

*Model.* In this paper it will be assumed that the earth's magnetic field is forced to lie in a cavity formed by the impinging solar plasma as shown in Figure 1 [F. S. Johnson, 1960a] The vertex of this surface occurs on the sunlit hemisphere near the earth-sun line at a distance of about 6 earth radii from the center

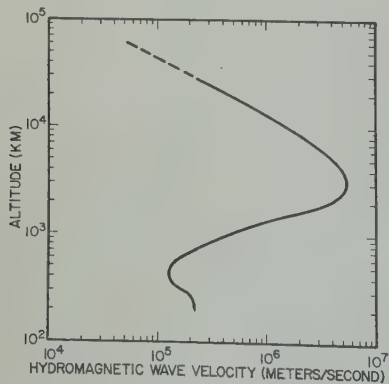


Fig. 2. Velocity of hydromagnetic waves in the geomagnetic equatorial plane vs. altitude for a surface field of 0.315 gauss. For other latitudes, the velocity shown should be multiplied by  $(1 + 3 \sin^2 \lambda)^{1/2}$ , where  $\lambda$  is the geomagnetic latitude. This curve applies for both longitudinal and transverse waves.

TABLE 1. Ion Number Density, Average Molecular Weight, and Hydromagnetic Velocity vs. Altitude above the Earth's Surface in the Magnetic Equatorial Plane

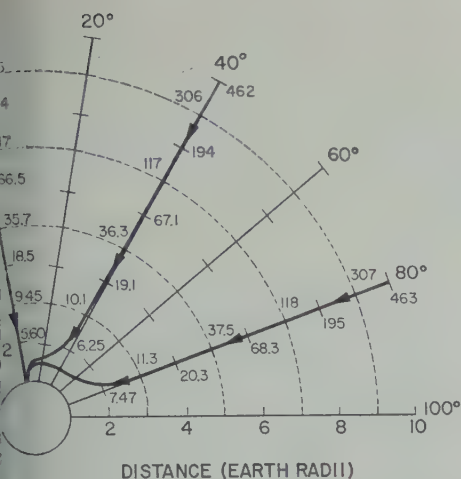
| Altitude, km | Ion Number Density, ions/meter <sup>3</sup> | Average Ionic Molecular Weight | Velocity, meter/sec |
|--------------|---|--------------------------------|---------------------|
| 200          | $3.31 \times 10^{11}$                       | 23.0                           | 2.27                |
| 250          | $3.72 \times 10^{11}$                       | 18.2                           | 2.36                |
| 300          | $6.76 \times 10^{11}$                       | 16.3                           | 1.80                |
| 400          | $1.10 \times 10^{12}$                       | 16.0                           | 1.37                |
| 500          | $1.00 \times 10^{12}$                       | 16.0                           | 1.40                |
| 600          | $6.30 \times 10^{11}$                       | 15.9                           | 1.66                |
| 800          | $2.06 \times 10^{11}$                       | 15.7                           | 2.69                |
| 1,000        | $6.34 \times 10^{10}$                       | 15.3                           | 4.51                |
| 1,500        | $6.67 \times 10^9$                          | 10.4                           | 1.38                |
| 2,000        | $2.56 \times 10^9$                          | 3.09                           | 3.41                |
| 2,500        | $2.04 \times 10^9$                          | 1.26                           | 5.05                |
| 3,000        | $1.75 \times 10^9$                          | 1.02                           | 5.18                |
| 4,000        | $1.28 \times 10^9$                          | 1.00                           | 4.44                |
| 5,000        | $1.01 \times 10^9$                          | 1.00                           | 3.80                |
| 7,000        | $8.12 \times 10^8$                          | 1.00                           | 2.61                |
| 10,000       | $6.51 \times 10^8$                          | 1.00                           | 1.58                |
| 15,000       | $5.12 \times 10^8$                          | 1.00                           | 8.05                |
| 25,000       | $4.05 \times 10^8$                          | 1.00                           | 2.89                |

of the earth. This distance is determined by the equilibrium condition

$$\rho v^2 \approx B^2/\mu_0$$

(mks units), where  $\rho$  and  $v$  are the solar mass density and velocity, respectively,  $\mu_0$  is the permeability of a vacuum, and  $B$  is the magnetic field strength just inside the solar wind. The solar wind will push closest to the earth along the earth-sun line. The radial distance from the earth to the boundary between the solar wind and the geomagnetic field will increase monotonically away from the earth-sun line, since, at large angles, the solar wind strikes the geomagnetic field obliquely and therefore with less pressure than at the subsolar point. An analytical expression for the shape of the surface is not available. For the purpose of calculation, it will be arbitrarily assumed that in the geomagnetic equatorial plane, the distance from the center of the earth to the boundary between the solar wind and the geomagnetic field increases smoothly from the subsolar point by a factor of 1.3 at points 60° either side. The existence of a steady, quiet





3. Three paths in the geomagnetic equatorial plane of hydromagnetic rays propagating to the earth from a distance of 10 earth radii. The transit times in seconds for each ray to the earth are given at each earth radii.

wind is not essential to this model. Since the velocity of the solar plasma is much greater than the hydromagnetic-wave velocity far from the earth, a surface such as that shown in Figure 2 can be established before any important hydromagnetic signals can propagate away from the earth.

Hydromagnetic waves will be generated over the broad front-surface by the impact of the solar plasma on the geomagnetic field. These waves carry the effects of the impact to the earth's surface, where they are observed as variations in the geomagnetic field. A sudden increase in the intensity of the solar wind results in the generation of a large-amplitude hydromagnetic wave over the entire front surface. The arrival of this wave at the earth's surface results in a world-wide sudden commencement.

The hydromagnetic velocity  $V$  is given by the relation

$$V = B/(\mu_0 \rho)^{1/2}$$

where  $B$  is the magnetic field strength and  $\rho$  is the plasma mass density. A geomagnetic disturbance will propagate with this velocity. The hydromagnetic-wave velocity, shown as a function of altitude in Figure 2, was computed assuming a centered dipole magnetic field and

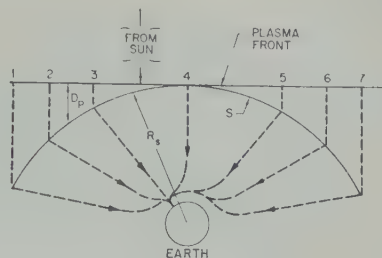


Fig. 4. The distance  $D_p$  traveled by various points on the impinging solar-plasma front to the surface limit of the geomagnetic field  $S$ , and hydromagnetic ray paths extending from the source  $S$  at a distance  $R_s$  from the center of the earth to a point on the earth  $40^\circ$  from the noon meridian, all on the plane of the geomagnetic equator.

utilizing the ion-density values given by *F. S. Johnson* [1960b] and the ion-composition observations of *C. Y. Johnson, Meadows, and Holmes* [1958]. This curve is similar to two previously reported curves presented by *Dessler* [1958] and by *Francis, Green, and Dessler* [1959]. The new curve is based on newer densities and is nearly the geometric mean of the previous curves. Table 1 gives the ion number density, the average ionic molecular weight, and the velocity of hydromagnetic waves as a function of altitude.

Using the new velocity profile, the ray paths and transit times for hydromagnetic waves propagating in the magnetic equatorial plane were calculated in accordance with Fermat's principle in the manner described by *Francis, Green, and Dessler* [1959]. The results are shown in Figure 3.

As was stated above, hydromagnetic waves are generated by the impinging solar plasma over a broad source at the surface limit of the geomagnetic field. The transit time of a hydromagnetic wave to a given point on the surface of the earth varies markedly with the position of its source on the surface of the geomagnetic field. Therefore, rays that leave the various regions of the field surface at the same time will, in general, arrive at a given point on the earth at different times. Furthermore, the impact of a plane plasma front on the field does not occur simultaneously along the curved geomagnetic surface but occurs first near the sun-earth line. Thus, hydromagnetic waves are not generated

TABLE 2. Transit Time of the Hydromagnetic Waves Generated across a Broad Source in the Earth's Geomagnetic Equatorial Plane by the Impact of a Sharp Front of Solar Plasma

We consider the seven ray paths shown in Figure 4.  $T_p$  is the time in seconds taken by the 1000-km plasma front to traverse the distance  $D_p$  (given in earth radii).  $T_{hm}$  is the time, in seconds, for a hydromagnetic ray generated at the surface  $S$ , a distance  $R_s$  (earth radii) from the earth's center, to propagate to a point on the earth  $40^\circ$  from the noon meridian.  $T_t$  is the total transit time in seconds, i.e., the sum of  $T_p$  and  $T_{hm}$ . The rise time is given as the difference between  $T_t(7)$  and  $T_t(4)$ , where the arguments refer to the numbered ray paths.  $\theta$  is the angular distance traveled in degrees measured clockwise by the hydromagnetic ray from earth meridian to source meridian.

| Ray No.<br>$\theta$                                    | 1<br>$-40^\circ$ | 2<br>$-20^\circ$ | 3<br>$0^\circ$ | *<br>$20^\circ$ | 4<br>$40^\circ$ | *<br>$60^\circ$ | 5<br>$80^\circ$ | 6<br>$100^\circ$ | 7<br>$120^\circ$ |
|--|------------------|------------------|----------------|-----------------|-----------------|-----------------|-----------------|------------------|------------------|
| Case I: $R_s \sim 4R_E$ Rise time $\sim 46$ seconds    |                  |                  |                |                 |                 |                 |                 |                  |                  |
| $D_p$  | 3.3              | 1.7              | 0.7            | 0.1             | 0.0             | 0.1             | 0.7             | 1.7              | 3.3              |
| $R_s$  | 5.2              | 4.7              | 4.3            | 4.1             | 4.0             | 4.1             | 4.3             | 4.7              | 5.2              |
| $T_p$  | 21.0             | 10.8             | 4.5            | 0.6             | 0.0             | 0.6             | 4.5             | 10.8             | 21.0             |
| $T_{hm}$   | 41.2             | 29.6             | 22.5           | 19.9            | 19.1            | 20.8            | 24.3            | 31.7             | 41.2             |
| $T_t$  | 62.2             | 40.4             | 27.0           | 20.5            | 19.1            | 21.4            | 28.8            | 42.5             | 62.2             |
| Case II: $R_s \sim 6R_E$ Rise time $\sim 142$ seconds  |                  |                  |                |                 |                 |                 |                 |                  |                  |
| $D_p$  | 4.7              | 2.5              | 1.1            | 0.2             | 0.0             | 0.2             | 1.1             | 2.5              | 4.7              |
| $R_s$  | 7.8              | 7.0              | 6.4            | 6.2             | 6.0             | 6.2             | 6.4             | 7.0              | 7.8              |
| $T_p$  | 30.0             | 16.0             | 7.0            | 1.2             | 0.0             | 1.2             | 7.0             | 16.0             | 30.0             |
| $T_{hm}$   | 177              | 117              | 83.8           | 75.8            | 67.1            | 76.6            | 85.8            | 119              | 177              |
| $T_t$  | 207              | 133              | 90.8           | 77.0            | 67.1            | 77.8            | 92.8            | 135              | 207              |
| Case III: $R_s \sim 8R_E$ Rise time $\sim 388$ seconds |                  |                  |                |                 |                 |                 |                 |                  |                  |
| $D_p$  | 6.2              | 3.4              | 1.5            | 0.3             | 0.0             | 0.3             | 1.5             | 3.4              | 6.2              |
| $R_s$  | 10.4             | 9.3              | 8.5            | 8.2             | 8.0             | 8.2             | 8.5             | 9.3              | 10.4             |
| $T_p$  | 39.5             | 21.7             | 9.6            | 1.9             | 0.0             | 1.9             | 9.6             | 21.7             | 39.5             |
| $T_{hm}$   | 540              | 347              | 245            | 213             | 194             | 214             | 247             | 349              | 542              |
| $T_t$  | 580              | 369              | 255            | 215             | 194             | 216             | 257             | 371              | 582              |

\* These rays are not shown in Figure 4.

simultaneously along the broad source. The above two effects produce a spread in the arrival times at an observing point on the earth of signals generated over the broad source. This delay causes a gradual build-up of the SC amplitude, giving a rise time of about 1 to 6 minutes.

Figure 4 depicts on the plane of the geomagnetic equator: the earth, the surface limit of the geomagnetic field  $S$ , and the impinging 1000-km/sec plasma front arriving from the sun. The plasma front first contacts the surface  $S$  at position 4, and later at positions 3 and 5, 2 and 6, and 1 and 7. Hydromagnetic rays from the surface  $S$ , which is a distance  $R_s$  from the center of the earth, travel in the indicated manner and arrive at a point on the earth arbitrarily chosen, for this figure, to lie on a meridian  $40^\circ$  from the noon meridian.

Table 2 gives the resulting transit times  $T_p$  and  $T_{hm}$  for individual rays. These transit times are, respectively, the time for the plasma front to travel a distance  $D_p$  to the surface  $S$ , and the time for a hydromagnetic wave to travel from the surface  $S$  to the earth (see Fig. 4). A measure of the SC rise time for this particular case is then obtained by subtracting the minimum total transit time  $T_t(4)$  from the maximum total transit time  $T_t(7)$ . As indicated by the three cases shown, the SC rise time is strongly dependent upon the source-earth distance  $R_s$ . It should be noted that, although it is possible to obtain greater total transit times by considering paths farther around on the extended source, the hydromagnetic attenuation suffered by these signals makes their contribution to the total SC rise time negligible.

The SC rise time is related to the distance

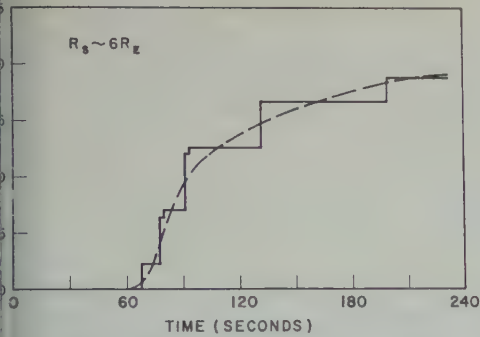


Fig. 5. Amplitude built up by hydromagnetic (separated by meridians of  $20^\circ$  angular distance) propagating in the geomagnetic equatorial plane to a point on the earth  $40^\circ$  from the noon meridian vs. the total transit time. The dashed curve was drawn so that the areas under the continuous and discontinuous curves are equal.

on the earth to the surface  $S$  as shown by the results given in Table 2. The SC amplitude would be larger when the distance  $R_s$  to the surface  $S$  is small because the plasma front must have been stronger to push  $S$  in to smaller  $R_s$ , and therefore presumably generated stronger hydromagnetic waves. Thus, a negative correlation between SC amplitude and rise time can be expected, that is, the SC rise time should show an inverse relationship to the SC amplitude.

The various waves that arrive at a given point on the earth undergo some degree of geometrical attenuation because they are spread out by reflection. The effect of this geometrical attenuation has been estimated by computing the divergence of rays from common points on the surface  $S$  (indicated in Fig. 4). Figure 5 shows the hydromagnetic signal amplitude built up in steps at the arrival at the earth of hydromagnetic signals from  $S$ , plotted against the corresponding total transit times (given in Table 2). The amplitude of the steps was computed by assuming continuity of hydromagnetic-wave flux through a thin slab containing the geomagnetic equator and by assuming uniformity of signal strength along the broad source  $S$ .

**Conclusions.** Assuming that the impact of a sharp solar-plasma front on the geomagnetic field generates a hydromagnetic wave, observed

at the earth's surface as an SC, the essential points made in this paper are: (1) because of the different transit times through the geomagnetic field, different parts of the hydromagnetic signal will arrive at a point on the earth at substantially different times; (2) this spread of arrival times constitutes the major portion of the observed SC rise time.

The results of the calculations of SC rise time are very sensitive to the shape of the boundary between the geomagnetic field and the solar wind. Since the shape of this boundary is not known either from experiments or from theory, the results presented here can be regarded only as illustrating a simple model whereby the sudden impact of a sharp solar-plasma front on the geomagnetic field may produce a relatively long SC rise time.

#### REFERENCES

- Chapman, S., and V. C. A. Ferraro, A new theory of magnetic storms, *Terrest. Magnetism and Atmospheric Elec.*, **36**, 77-97 and 171-186, 1931.
- Chapman, S., and V. C. A. Ferraro, The theory of the first phase of a geomagnetic storm, *Terrest. Magnetism and Atmospheric Elec.*, **45**, 245-268, 1940.
- Dessler, A. J., The propagation velocity of worldwide sudden commencements of magnetic storms, *J. Geophys. Research*, **63**, 405-408, 1958.
- Dessler, A. J., Ionospheric heating by hydromagnetic waves, *J. Geophys. Research*, **64**, 397-401, 1959.
- Dessler, A. J., and E. N. Parker, Hydromagnetic theory of geomagnetic storms, *J. Geophys. Research*, **64**, 2239-2252, 1959.
- Francis, W. E., M. I. Green, and A. J. Dessler, Hydromagnetic propagation of sudden commencements of magnetic storms, *J. Geophys. Research*, **64**, 1643-1645, 1959.
- Francis, W. E., A. J. Dessler, and R. Karplus, Attenuation of hydromagnetic waves in the ionosphere, *Bull. Am. Phys. Soc.*, [2] **5**(4), 316, 1960.
- Hines, C. O., On the geomagnetic storm effect, *J. Geophys. Research*, **62**, 491-492, 1957.
- Johnson, C. Y., E. B. Meadows, and J. C. Holmes, Ion composition of the arctic ionosphere, *J. Geophys. Research*, **63**, 443-444, 1958.
- Johnson, F. S., The composition of outer space, *Astronautics*, **54**, 30-31, 1960a.
- Johnson, F. S., The ion distribution above the  $F_2$  maximum, *J. Geophys. Research*, **65**, 577-584, 1960b.

(Manuscript received June 3, 1960.)





# A Note on Harmonic Analysis of Geophysical Data with Special Reference to the Analysis of Geomagnetic Storms

MASAHISA SUGIURA

*Geophysical Institute, University of Alaska  
College, Alaska*

**Abstract.** Some geophysical characteristics tend to have a fixed distribution relative to the sun. An example is the distribution of air temperature on an ideal earth that is perfectly symmetrical (e.g., in its pattern of land and water) about its axis of rotation. In such a case the geophysical characteristic at any fixed station on the earth undergoes a daily variation that depends only on local time (and latitude and season). This simple pattern of daily change may be modified by intrinsic changes in the solar influences on the earth. The harmonic components of the daily variation at any station may then undergo phase changes, causing frequency shifts, in some respects corresponding to Doppler shifts of frequency in optical or sonic phenomena. Care is then needed if the results of harmonic analysis are to be properly interpreted. Such interpretation is discussed with reference to the parts Dst and DS of the magnetic storm variations. Like caution must be observed if the amplitude of a harmonic variation changes with fixed phase.

**Introduction.** In geophysics one frequently has phenomena periodic both in space and in time. Because of the rotation of the earth about its axis, temporal coordinates—local time, universal time, or time defined in a special manner with reference to some event—cannot be the spatial coordinates that specify the positions of observatories relative to the sun, and, in general, uniquely related to each other. In analyzing observational records the distinction between the temporal and the geometrical coordinates must be clearly borne in mind, although they are often intertransformable.

In particular, progressive change of a phenomenon with local time as observed at a fixed observatory, and the spatial distribution at the instant of time of the same phenomenon at different longitudes, relative to the sun, should not be confused. These two variations may be regarded as equivalent only when the phenomenon concerned stays stationary relative to the sun.

In analyzing a set of data harmonically, one must ask: *with respect to what coordinate is the variation analyzed?*

**1. Harmonic analysis of a variation with constant amplitude and changing phase.** Let us suppose that a wave is stationary in the reference frame  $S$  that is at rest relative to the sun,

and that in  $S$  the wave is described by  $c \sin m\lambda$ , where  $c$  is constant and  $m$  an integer; the variable  $\lambda$  is reckoned in angular measure. The reference frame  $S_1$  is traveling with a constant (angular) velocity  $\omega_1$  in the direction of increasing  $\lambda$ . It is assumed that for motions considered in this paper the Galilean transformation of coordinates is valid.

If the origins of  $S$  and  $S_1$  coincide at time  $t = 0$ ,  $\lambda$  and  $\lambda_1$  are connected by  $\lambda = \lambda_1 + \omega_1 t$ , where  $t$  is 'absolute' time. To the observer  $A$ , who is at rest at the origin of  $S_1$ , the wave appears as  $c \sin m\omega_1 t$ . If he reckons time in angular measure, according to  $t' = \omega_1 t$ , his description of the wave is  $c \sin mt'$  in terms of his (local) time  $t'$ , or  $c \sin m\lambda_1$  in terms of his longitude  $\lambda_1$  relative to the sun.

An idealized quiet daily variation ( $S_q$ ) that changes neither its amplitude nor its phase may be regarded as an example of such a wave.

Next we consider a wave traveling relative to  $S$  with a constant (angular) velocity  $\omega_0$  in the direction of increasing  $\lambda$ . In  $S$ , the wave is  $c \sin m(\lambda - \omega_0 t)$ ; see Figure 1. In  $S_1$ , which is moving with velocity  $\omega_1$ , the wave is

$$c \sin m[\lambda_1 + (\omega_1 - \omega_0)t]$$

The observer  $A$ , who is at rest at the origin of  $S_1$ , therefore, observes the wave to be

$$c \sin m(1 - \omega_0/\omega_1)\omega_1 t$$

In terms of  $t'$ , or  $\lambda_A$ , this latter is

$$c \sin m^* t' \quad \text{or} \quad c \sin m^* \lambda_A$$

where

$$m^* = m(1 - \omega_0/\omega_1) \quad (1)$$

The frequency  $m^*$  is the familiar (classical) Doppler-shifted frequency.

The DS variation of magnetic storms is an example of this kind of wave except that its phase changes at a varying rate.

Without knowing the changing phase (with respect to  $S$ ), which the motion of the wave amounts to, the observer  $A$  analyzes his observational data of the variation  $c \sin m^* \lambda_A$  in the interval  $\lambda_A = -\pi$  to  $\pi$  as if the wave were stationary to  $S$ . He will, therefore, obtain *not* the  $m$ th harmonic alone, but, in general, all frequencies from 0 to  $\infty$ , the spectrum being dependent on  $\omega_0/\omega_1$ .

Expanding the function  $c \sin m^* \lambda$  in a Fourier sine series in the interval  $\lambda = -\pi$  to  $\pi$ , we obtain

$$c \sin m^* \lambda = c \sum_{k=1}^{\infty} b_k \sin k\lambda$$

where

$$b_k = (-1)^k (2k \sin m^* \pi) / \{\pi(m^{*2} - k^2)\} \quad (2)$$

if  $m^*$  is not an integer; and

$$b_k = \delta_{km^*} \quad (2')$$

if  $m^*$  is an integer, where  $\delta_{ij} = 0$  when  $i \neq j$  and  $\delta_{ii} = 1$  when  $i = j$ .

When  $m^*$  deviates only by a small amount, say  $\epsilon$ , from an integer, say  $n$ , so that  $m^* = n + \epsilon$  ( $\epsilon \ll n$ ), then from (2) we can readily show that  $b_n$ , given by (2), tends to 1 as  $\epsilon$  tends to 0. This corresponds to a complete shift of frequency from  $m$  to  $n$ ; in this case other harmonics have zero amplitude.

Figure 2 shows the coefficients  $b_k$  as functions of  $m^*$ , for  $k = 1, 2, 3$ , and 4. For each value of  $k$ ,  $b_k$  is 1 at  $m^* = k$ , and is maximum at  $m^*$  rather less than  $k$ ; at this value of  $m^*$ , which satisfies

$$\cot m^* \pi = 2m^* / \{\pi(m^{*2} - k^2)\}$$

$b_k$  is

$$2k\{\pi^2(m^{*2} - k^2)^2 + 4m^{*2}\}^{-1/2}$$

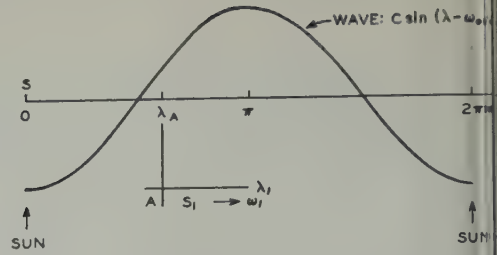


Fig. 1. The wave travels with a constant velocity  $\omega_0$  relative to  $S$ . The wave appears in  $S$  as

$$c \sin (\lambda - \omega_0 t)$$

In  $S_1$ , traveling with a constant velocity  $\omega_1$ , wave is

$$c \sin \{\lambda_1 + (\omega_1 - \omega_0)t\}$$

The observer  $A$  at the origin of  $S_1$  sees

$$c \sin (1 - \omega_0/\omega_1)\omega_1 t$$

or

$$c \sin (1 - \omega_0/\omega_1)\lambda_A$$

which is very nearly

$$1 + (\delta/k) + (1 - k^2\pi^2/2) (\delta/k)^2$$

where  $\delta = k - m^*$ .

This result has an important implication in the interpretation of harmonic analysis. Take, for instance, the case when  $m^* = 1.4$ . For this value of  $m^*$ ,  $b_1$  and  $b_2$  are nearly equal (Figure 2). Thus the observer  $A$  obtains the first and second harmonics that are nearly of equal magnitude, although the wave is a single harmonic in  $S$ . If  $m = 1$ ,  $\omega_0/\omega_1 = -0.4$  gives  $m^* = 1.4$ ; if  $m = 2$ ,  $\omega_0/\omega_1 = 0.3$  gives  $m^* = 1.4$ , etc. From the harmonic analysis alone, there is no way of discriminating between these possibilities: either the wave is not a single harmonic in  $S$ , or if the velocity  $\omega_0$  is not constant in time, the situation is even more complex.

This example illustrates the importance of a careful examination of the treatment of data and the interpretation of the result of harmonic analysis or of the synoptic representation of data.

It is further pointed out in connection with magnetic storm analysis that, in general, the variation to be expanded in the interval  $-\pi$  will be of the form



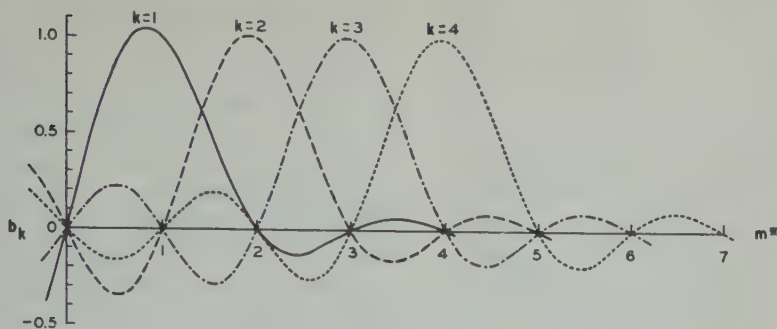


Fig. 2. The coefficients  $b_k$  as functions of  $m^*$ , for  $k = 1, 2, 3$ , and  $4$ .

$$\{a_m \cos m(1 - \omega_0/\omega_1)\lambda$$

$$b_m \sin m(1 - \omega_0/\omega_1)\lambda\}$$

In this case, because of the cosine terms, Dst (see section 4) will not be simply  $a_0$ , but

$$a_0 + \sum_{m=1}^{\infty} \frac{a_m \sin m(1 - \omega_0/\omega_1)\pi}{m\pi(1 - \omega_0/\omega_1)}$$

Therefore, in determining Dst from the observed data, some precaution must be taken to eliminate the influence of the Doppler shift involved in

*Harmonic analysis of a variation with changing amplitude and constant phase.* Another example, of somewhat different nature, is the harmonic analysis of a variation, which is a simple harmonic in  $\lambda$ , in  $S_1$  at any instant of time, but which changes its amplitude with time.

Suppose that such a wave is expressed by

$$c(t) \sin m\lambda,$$

where  $c(t)$  is a function of  $t$ . With respect to  $S_1$ , the wave is

$$c(t) \sin m(\lambda_1 + \omega_1 t)$$

If the observer  $A$  uses this coordinate  $\lambda_A$  as the time coordinate, he will describe the wave as

$$c'(\lambda_A) \sin m\lambda_A$$

where

$$c'(\lambda_A) \equiv c(\lambda_A/\omega_1)$$

If he analyzes this variation harmonically in the interval  $\lambda_A = -\pi$  to  $\pi$ , he will, in general,

obtain harmonics of harmonic numbers from 0 to  $\infty$ .

If  $c'(\lambda)$  is expressible by a power series of  $\lambda$ , i.e.

$$c'(\lambda) = \sum_{l=0}^{\infty} c_l \lambda^l$$

the integrals giving the coefficients of the Fourier expansion of  $c'(\lambda) \sin m\lambda$  can readily be integrated by repeated application of integration by parts.

For an example, if  $c'(\lambda)$  is expressed as

$$c'(\lambda) = c_0 + c_1 \lambda + c_2 \lambda^2 \quad (3)$$

the coefficients of the expansion

$$c'(\lambda) \sin m\lambda = a_0 + \sum_{k=1}^{\infty} (a_k \cos k\lambda + b_k \sin k\lambda) \quad (4)$$

are given by

$$a_0 = (-1)^{m+1}(c_1/m) \quad (5)$$

and for  $k > 0$ ,  $k \neq m$ ,

$$a_k = (-1)^{k+m+1}(2mc_1)/(m^2 - k^2) \quad (6)$$

$$b_k = (-1)^{k+m}(8kmc_2)/(m^2 - k^2)^2 \quad (7)$$

and for  $k = m$ ,

$$a_m = -c_1/(2m) \quad (8)$$

$$b_m = c_0 + \{\pi^2/3 - 1/(2m^2)\}c_2 \quad (9)$$

The effect is unlikely to be important in the analysis of the solar or lunisolar, daily geomagnetic or ionospheric variations, in which the day-to-day change is not so great. Even at the geomagnetic equator, where the day-to-day

change in  $S_e$  is appreciable, the effect is only slight. However, when variations with rapidly changing amplitude are analyzed, this problem may be worth considering.

4. *DS variation in magnetic storms.* Chapman [1952] and Sugiura and Chapman [1956-1958; 1960] analyzed the magnetic storm field in two parts, Dst and DS. In our conception the separation into the two parts was made for each instant of storm time, the storm field being regarded as a function of longitude relative to the sun and of geomagnetic latitude. Dst was defined as the mean over all longitudes, and DS the deviation from it.

It was assumed that, in each latitude belt, variations observed at different stations and for different storms are all equivalent.

The method used by us for the analysis of DS for each of the first 6 hours (section 14 of our 1960 paper) expresses the above 'geometrical' concept of DS. For the subsequent phases of the storm, DS was determined for each of 6- or 8-hour intervals.

It was found that both the phase and the amplitude of DS vary very rapidly during the first several hours, and that these changes become slower as the storm progresses.

Though the concept of Dst and DS was clear in our mind, the question of the 'Doppler effect' was not recognized when the analysis was made.

The fact of the appearance and nonappearance of the Doppler effect in DS according to the method of analysis used may be illustrated in the following way.

Suppose that a wave is  $c \sin m(\lambda - \omega_0 t)$  in the reference frame  $S$  that is stationary with regard to the sun. With reference to the frame  $S_1$  that is moving with velocity  $\omega_1$ , the wave is  $c \sin m[\lambda_1 + (\omega_1 - \omega_0)t]$ . If this wave is observed at an instant  $t$ , for all  $\lambda_1$ , it varies, apart from some constant phase angle, as  $c \sin m\lambda_1$  in space, and hence there is no Doppler effect. If the wave is seen by an observer at  $\lambda_1 = \lambda_0$ , he observes the wave to vary, apart from some phase angle, as  $c \sin m(\omega_1 - \omega_0)t$  in time. Hence, the frequency is Doppler-shifted.

If he observes the (time) rate of change, instead of the change itself, he is *always* observing the rate of change at Doppler-shifted frequencies.

With these remarks in mind one can draw the

following conclusions regarding the actual treatment of the data.

1. If DS is determined for any instant of storm time, the Doppler effect does not appear. In this type of analysis each observatory contributes only one instantaneous value per storm.

2. If DS is derived from data for any interval of time, regardless of the length of interval, the question of Doppler effect arises. In this case each observatory contributes two or more values per storm, depending on the length of interval.

3. If DS is determined from the rate of change, such as hour-to-hour differences, the Doppler effect *inevitably* appears, even if the rate of change refers to some instant of time. In fact, this is the limiting case of paragraph 2 when the interval is made infinitely short.

In our analysis the rate of change was used to determine DS for 6- or 8-hour intervals. For the first 6 hours, DS was also determined from individual hours, likewise using the rate of change. Though the phase (with respect to the reference frame fixed to the sun) changes rapidly during these hours, the rate of change may be considered approximately constant in each 1-hour interval for which the rate of change (hour-to-hour difference) was computed. For the subsequent intervals the rate of change in the phase may be considered nearly constant within each interval. Therefore, the harmonic coefficients of DS for Dst for the first 6 individual hours and for the subsequent intervals can, in principle, be interpreted in a definite if not simple manner in terms of a set of parameters.

Yokouchi [1958] determined DS (and Dst) for overlapping intervals of 24 hours and for individual storm hours, using instantaneous hourly values, measured from prestorm values, for many storms for the Kakioka Magnetic Observatory. He showed some important features that had not been found by Chapman and Sugiura because of the difference in the method of analysis. Professor Chapman brought this discrepancy between Yokouchi's and our results to my attention. In the process of studying the difference in the treatment of data, the present paper was written as a by-product. A preliminary study has confirmed Yokouchi's results and extended them further. A more detailed study is now being made.

In conclusion it is emphasized that due caution should be taken in data analysis or graphical representation when the variation dealt with involves simultaneously both time and space coordinates as variables.

*Acknowledgments.* I wish to thank Professor S. Chapman for valuable discussions and for reading the manuscript before publication. I am indebted to Mrs. Sandra J. Fuller for her assistance in the numerical computation.

The research reported in this paper has been sponsored by the Geophysics Research Directorate, Air Force Cambridge Research Center, Massachusetts, through contract AF 19(604)-2163.

## REFERENCES

Chapman, S., The morphology of geomagnetic

storms: an extension of the analysis of DS, the disturbance local-time inequality, *Ann. geophys. Rome*, 5, 1-19, 1952.

Sugiura, M., and S. Chapman, Morphology of magnetic storms, *Geophys. Inst., College, Alaska, Final Repts.*: weak storms, AF 19(604)1048, p. 126, 1956; moderate storms, AF 19(604)1732, p. 54, 1957; great storms, AF 19(604)2163, p. 60, 1958.

Sugiura, M., and S. Chapman, The average morphology of geomagnetic storms with sudden commencement, *Abhandl. Akad. Wiss. Göttingen, Math.-physik. Kl., Sonderheft 4*, 1960.

Yokouchi, Y., Development of SD-field with storm-time at Kakioka, *Mem. Kakioka Magnetic Observatory*, 8, 13-32, 1958.

(Manuscript received June 17, 1960.)





# Direct Measurement of Particles Producing Visible Auroras

CARL E. McILWAIN

*Department of Physics and Astronomy  
State University of Iowa  
Iowa City, Iowa*

**Abstract.** The nature of the particles producing two visible auroras at Fort Churchill, Canada, has been determined with the aid of rocket-borne detectors. Magnetic fields were used to separate the proton and electron components of the particle flux. It was found in each case that a major fraction of the auroral light was produced by electrons with energies of less than 10 keV.

IGY rocket II6.26F was fired into a quiescent auroral glow of about intensity I. The integral number energy spectrum of protons incident upon the atmosphere during this flight was measured to be approximately  $2.5 \times 10^6 \exp(-E/30)$  protons/sec  $\text{cm}^2$  ster over the range 80 to 250 keV, where  $E$  is the proton energy in keV. The observed altitude dependence of the electron energy flux could have been produced by an integral number energy spectrum equal to  $2.5 \times 10^6 \exp(-E/5)$  electrons/sec  $\text{cm}^2$  ster over the range 3 to 30 keV.

IGY rocket II6.27F was fired into a bright active auroral arc. At least 75 per cent of the light in this aurora was produced by nearly monoenergetic electrons with about 6-keV energy. The electron flux varied rapidly with time; peak values were about  $5 \times 10^{10}$  electrons/sec  $\text{cm}^2$  ster. The presence of monoenergetic electrons strongly suggests an electrostatic acceleration mechanism.

## INTRODUCTION

Extensive observations of auroras have been made with ground-based optical instruments [Frey, 1955]. These investigations have provided excellent descriptions of most auroral phenomena. In particular, it has been established that the auroral light is due to energetic charged particles entering the earth's atmosphere (from alignment of the auroral features with the magnetic field), and that some of the particles are protons (from the presence of Doppler-broadened  $H_\alpha$  and  $H_\beta$  spectral lines). Other important results have been to determine the altitude distribution of the auroral light emission; to establish the frequency of occurrence at different places on the earth; and to show the correlation with magnetic disturbances and activity on the sun.

Rocket techniques greatly extend the kind of measurements that can be made. They are capable of providing such information about the auroral particles as: (1) the relative number of electrons and protons; (2) the intensities of the particle fluxes; (3) the energy spectra of the particles; (4) the effects of the particles, such as the production of light and ionization.

Information like this may eventually shed light on where the particles come from and how they are accelerated.

The present paper is based on the results obtained from two similar instrumentation packages that were placed on Nike-Cajun rockets and fired into visible auroras at Fort Churchill, Canada, during February 1958.

The first flight was into a faint aurora which contained the  $H_\beta$  line in its optical spectrum. An electron energy flux of about 20 ergs/sec  $\text{cm}^2$  was measured by the rocket instruments. The proton detector in the rocket indicated the incidence of approximately 1 proton for every 1000 electrons. The second flight was into a bright auroral arc. The rocket-borne detectors suggest that the total energy flux was as high as 2000 ergs/sec  $\text{cm}^2$  but the proton flux appeared to be less than on the previous flight.

## PURPOSE

This exploratory investigation was intended to obtain a gross characterization of the particles that produce visible auroras. Accuracy and resolution were intentionally sacrificed in order to provide sensitivity to particles over a wide

range of intensities and energy. This choice was fortunate in view of the large time and spatial variations observed and the fact that many of the estimates of particle energies and intensities based on optical measurements were in error by more than an order of magnitude.

### APPARATUS

The rocket instrumentation contained three charged-particle detectors, a photometer, and a magnetometer (see Figs. 1 and 2).

*Proton detector.* The instrument designed to detect protons consists of a disk of CsI (TI) scintillator 94 mg/cm<sup>2</sup> thick cemented to a photomultiplier tube (see Fig. 3). A 400-gauss magnetic field at the center aperture prevents electrons with less than 1-Mev energy from reaching the scintillator. The geometrical factor determined by the apertures is  $3.5 \times 10^{-3}$  cm<sup>2</sup> ster. About 40 micrograms/cm<sup>2</sup> of aluminum is evaporated on the crystal, reducing the light sensitivity by a factor of about 10<sup>4</sup>. The aluminized crystal on the face of the RCA type 6199 photomultiplier tube is pictured in Figure 4. The detector is sensitive to all kinds of energetic ions, but for convenience it will be referred to as the 'proton detector.' In particular, the instrument is not capable of distinguishing  $\alpha$  particles from protons.

The number of pulses from the photomultiplier tube corresponding to energy losses greater than 25 kev (45-kev protons before energy loss in the aluminum) are counted. In addition, the pulse height distribution corresponding to energy losses between 25 and 2500 kev is obtained by the following procedure:

1. In the rocket the pulse heights are converted into pulse lengths. The conversion process is inhibited for 10<sup>-2</sup> second after the start of each conversion to prevent overlap and to permit input rates of up to 10<sup>4</sup>/sec without distorting the distribution.

2. On the ground the pulse lengths are used to gate a stable oscillator, thus providing a digital representation of each pulse height.

3. The above procedure is identical in principle to the operation of the analog to digital converter in the 256-channel pulse height analyzer manufactured by Radiation Counter Laboratories. It is therefore possible to feed the



APERTURES AND  
MAGNETIC FIELDS

GEIGER TUBE, PROTON  
AND ELECTRON DETECTOR

PHOTOMETER

MAGNETOMETER

Fig. 1. Photograph of the rocket instrumentation, showing the placement of the detectors. The height of the instrumentation is 94 cm.

signals derived from the rocket telemetry with appropriate command signals into an analyzer to obtain the pulse height distribution at the detector output.

The pulse height distribution width, with the system using a pulse generator, is about 10% over the range of 25 to 500 kev. The energy resolution of the detector is approximately 10% per cent at 100-kev energy loss. The relative pulse height calibration of the instruments was determined using a pulse generator. The absolute calibration was determined with Po<sup>210</sup>  $\alpha$  particles and Co<sup>57</sup>  $\gamma$  rays.

*Electron detector.* Figure 3 shows a section of the electron detector. The magnetic field of the 250-turn coil focuses electrons through the ring-shaped aperture onto the ring of (TI) scintillator if the energy of the elec-



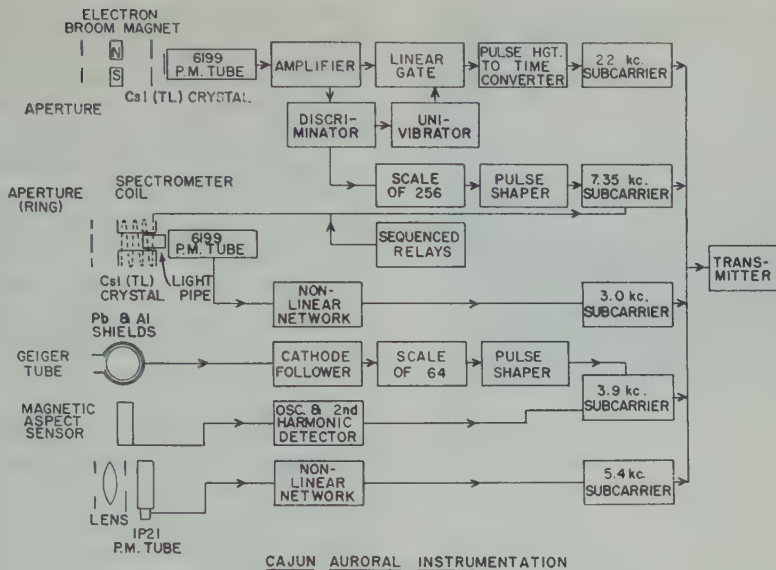


Fig. 2. Block diagram of the rocket instrumentation.

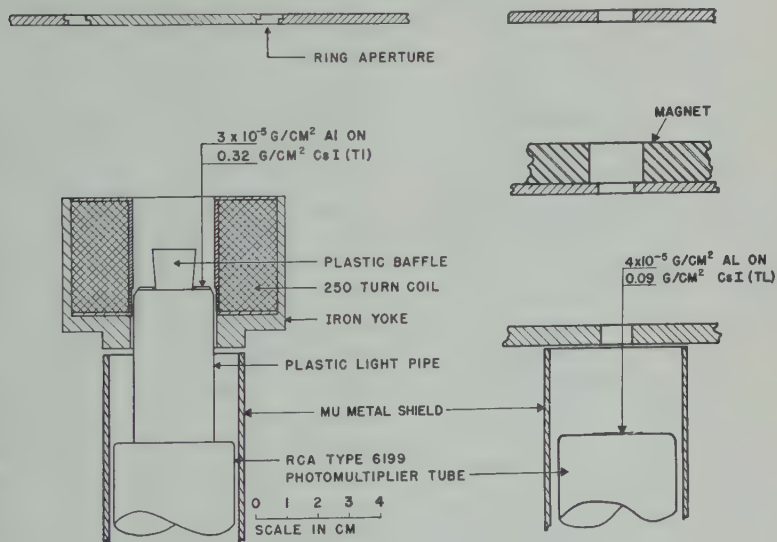


Fig. 3. Cross section of the electron (left) and proton (right) detectors, showing the geometry of the apertures and magnetic fields.

within about a factor of 1.7 of an optimum energy set by the current in the coil. Since the photomultiplier tube pulses caused by very low energy electrons are too small to be counted individually, they are integrated, thus producing a current proportional to the average energy per second in the crystal. This current is passed through a nonlinear network which pro-

duces a voltage approximately proportional to the logarithm of the energy flux over the range  $10^{-3}$  to  $10^2$  ergs/sec  $\text{cm}^2$  ster.

The crystal is covered by  $30 \mu\text{g}/\text{cm}^2$  of aluminum to reduce the light sensitivity. This aluminum film makes electrons with less than about 3-kev energy undetectable. The coil current is varied between 0 and 15 amperes by means of



Fig. 4. A photograph of an aluminized disk of CsI crystal cemented to the face of a photo-multiplier tube. The 1-cm-diameter crystal is the sensitive element in the proton detector.

sequenced relays in seven steps every 6 seconds. The highest current focuses electrons with energies up to 100 kev.

For this exploratory work, an intermediate baffle is omitted, therefore permitting rectilinear paths. This arrangement partially destroys the differential character of the instrument, but it has the advantage that the output with zero coil current is a measure of the total energy flux of all particles after passing through the  $30 \mu\text{g}/\text{cm}^2$  of aluminum covering the crystal. The geometrical factor for such rectilinear paths is  $8.5 \times 10^{-3} \text{ cm}^2 \text{ ster}$ . The geometrical factor for focused electrons is about twice this value.

The sensitivity of each crystal and photo-multiplier tube combination was obtained by measuring the output as a function of the distance from a calibrated  $\text{Po}^{210}$   $\alpha$ -particle source. By performing this measurement before and after aluminization, the thickness of the aluminum was also determined. Unfortunately, an isotropic monoenergetic electron source cannot be easily obtained to provide an accurate calibration of the energy selectivity of the instrument. The energy best focused by the coil for different coil currents was obtained by rotating the instrument in two planes in a parallel monoenergetic electron beam.

Protons capable of penetrating the aluminum on the crystal are not appreciably deflected by the magnetic field; therefore, the changes in the detector output from the value at zero coil current can be due only to electrons.

Both the proton detector and the electron

detector were aligned with the rocket axis. During the initial part of flight in the dense atmosphere an aerodynamic cone protected the detectors from damage. The cone was ejected a minute after launch. A  $\text{Po}^{210}$   $\alpha$ -particle source was placed on this cone to prove that it was ejected. A similar source was placed near the middle aperture of the proton detector. The pulse heights produced in the proton detector by these two sources provided a measure of atmospheric density in the altitude range 10 to 25 km during ascent, and the source near the middle aperture provided a measure from 35 to 25 km during descent.

**Geiger tube.** Also included in the rocket instrumentation is a Victoreen type 1B85 Geiger counter, shielded by 1/16 inch of lead except for one side next to a  $0.26 \text{ g}/\text{cm}^2$  aluminum window in the side of the instrumentation package. The geometrical factor for 1.0 to 5.0 Mev electrons in this configuration is  $20 \text{ cm}^2 \text{ ster}$ . For lower energy electrons the detector has low efficiency by way of X-ray production. The omnidirectional flux in electrons/sec  $\text{cm}^2 \text{ ster}$  required to produce 1 count/sec in this manner was measured to be approximately  $2 \times 10^{10}$  at 100 kev,  $1.7 \times 10^9$  at 40 kev,  $6 \times 10^8$  at 30 kev,  $6 \times 10^7$  at 20 kev, and  $10^{12}$  at 10 kev.

**Rocket photometer.** A photometer in the rocket instrumentation is used to measure the total directional intensity of visible auroral light. The photometer consists of an RCA 1P21 photomultiplier tube, apertures, a lens, and a Vycor window. The field of view is  $3.5^\circ$  by  $5^\circ$  directed perpendicular to the rocket axis. The photomultiplier tube current output is passed through a nonlinear network which provides a uniform sensitivity over the range  $10^{-8}$  to  $4.0 \text{ ergs}/\text{sec cm}^2 \text{ ster}$ .

**Ground photometer.** A photometer with a  $3.5^\circ$  by  $7^\circ$  field of view directed near the vertical is used to help determine the best time to fire the rockets and to provide a measure of auroral intensity during the rocket flights. Spectrometers are not used with either the rocket or ground rocket photometer. The photometers simultaneously weight the auroral spectrum according to the S-11 photocathode spectral sensitivity curve. The absolute calibration of this photometer was obtained using a secondary standard low brightness source obtained from the Defence Research Northern Laboratory at Fort Churchill, Canada.

**Aspect sensor.** A flux gate magnetometer manufactured by the Schonstedt Engineering Company is used to determine the aspect of the rocket with respect to the magnetic field. The sensing element of the magnetometer is tilted  $9^\circ$  from the rocket axis to improve the accuracy of aspect determination when the rocket is close to alignment with the magnetic field. The magnetic field at Fort Churchill is inclined  $6.3^\circ$  from the vertical; therefore, for any purposes the angle measured by the magnetometer can be considered to be the angle to the vertical.

#### ALTITUDE DETERMINATION

No radar or Doppler type of transponder is included in the instrumentation, in order to keep the total weight below 50 pounds. During the ascent, vents in the side of the outer instrumentation tube maintain the air density at the detectors near the ambient atmospheric density; therefore, the energy loss of the  $\alpha$  particles from the  $\text{Po}^{210}$  sources placed above the proton detector provides a direct measure of the rocket altitude. During descent the absence of the aerodynamic cone causes the air density at the detectors to be much lower than the ambient atmospheric density. This 'ram' density is approximately equal to  $0.9 \rho (M^2 + 0.36)$ , where  $\rho$  is the atmospheric density and  $M$  the Mach number [Havens, Holl, and LaGow, 1952].

An IBM 650 computer is used to obtain the complete rocket trajectory by starting with the measured time and altitude at second-stage burnout and trying different burnout velocities until an accurate fit to the measured ram density versus time during descent is obtained. The accuracy of this method is estimated to be better than  $\pm 1$  km. Stereoscopic photographs taken at the launch site and at a site 10 km to the west are used to determine the horizontal component of the rocket velocity at the time of second-stage burnout. To determine the horizontal position of the rocket during the remainder of the flight, this velocity is assumed to be constant.

#### PERFORMANCE

The instruments on both rocket flights performed as anticipated. The rocket antenna radiation pattern was nearly omnidirectional;

therefore, no information was lost because of poor antenna orientation. The transmitter frequency on the second flight was somewhat unstable, however, and caused some loss of information.

The Nike-Cajun rocket has been used to carry payloads of similar weight to altitudes greater than 160 km. In these two rocket flights the peak altitudes were approximately 120 km. As the altitude and velocity at second-stage ignition were normal in both, the low performance was presumably due to low second-stage thrust.

The low peak altitudes achieved resulted in an appreciable degradation of the data, owing to the fact that the atmospheric pressure is not known well enough at these altitudes to correct for the air absorption accurately and to the impossibility of detecting the very-low-energy particles that are stopped or degraded in energy by the air remaining above the rocket.

#### EVENTS DURING FEBRUARY 1958

A number of unusual geophysical events occurred during February 1958. A sudden commencement occurred at 0125 UT February 11 and was followed by a severe magnetic storm. A Forbush-type decrease [Winckler, Peterson, Hoffman, and Arnoldy, 1959; Lockwood, 1960] in the cosmic-ray intensity began at about 0200 UT February 11; by 066 UT sea-level neutron monitors registered 4 to 5 per cent decreases. On the night of February 11-12 a bright red aurora was observed to extend over a large range of latitudes and longitudes. The following two weeks were characterized by moderate magnetic activity and frequent occurrence of high-latitude auroras. The ground photometer records for the period February 10 through 25 were analyzed to obtain the percentage of time the auroral light intensity overhead at Fort Churchill was greater than various values. Excellent weather conditions permitted 80 hours of cloud-free observations to be obtained. Figure 5 shows that the night sky was never fainter than about  $3 \times 10^{-2}$  erg/sec  $\text{cm}^2$  (column) during this period and that the occurrence of higher intensities decreased approximately as the inverse square of the intensity. It also shows that for only about 10 minutes during this 80-hour period was



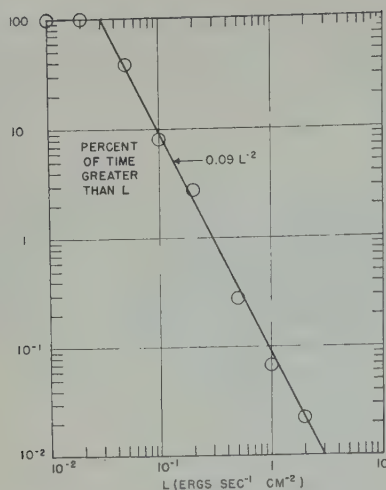


Fig. 5. Percentage of time the total visible overhead light flux (per  $\text{cm}^2$  column) was greater than  $L$  during 80 hours of observation between February 10 and February 25, 1958, at Fort Churchill, Canada.

the auroral intensity greater than the maximum intensity that occurred during the flight of rocket II6.27F.

#### RESULTS FROM IGY ROCKET II6.26F

*Characteristics of the aurora.* IGY rocket II6.26F was fired at 05 hr 34 min 31 sec UT on February 22, 1958, into a faint quiescent auroral glow. As can be seen in Figure 6, which is a photograph taken by a high-speed all-sky camera at Fort Churchill when the rocket was at peak altitude, the auroral light intensity was fairly constant over a large area. The lower curve in Figure 7 shows the light intensity measured by the ground photometer during this flight. The average value is about 0.05 erg/sec  $\text{cm}^2$  (col.) above the average night airglow value of 0.03 erg/sec  $\text{cm}^2$  (col.). The night airglow intensity at lower latitudes is usually less than 0.01 erg/sec  $\text{cm}^2$  (col.). The origin of this rather high night airglow intensity is unknown, but it may be the same as that of the light emissions that are labeled auroras because of their intensity and their time and spatial dependence. The absolute calibration of the ground photometer was roughly checked by visually noting that the 'night airglow' partly obscured the Milky Way.

On account of the diffuse nature of the auroral light emission, the altitude of the lower boundary of the aurora during this rocket flight could not be measured from the ground.

*Hydrogen spectrum.* The most important ground observations made during this flight were measurements by Montalbetti [1959] of the spectral line intensity in the vertical direction at Fort Churchill (20 km from the rocket launch site). The spectrometer employed had an resolution and scanned the region from 4580 to 4950 Å in 12 seconds every 15 seconds. The average  $H_\beta$  intensity while the rocket was at 100 km was about  $6 \times 10^7$  quanta/sec  $\text{cm}^2$  (col.).

*Protons.* The flux of particles versus altitude recorded by the rocket proton detector is shown in Figure 8. The Geiger tube registered less than 0.2 particle/sec  $\text{cm}^2$  ster with regard to cosmic-ray flux; therefore the particles detected must have been protons,  $\alpha$  particles, or heavy ions. It will be assumed that the particles were protons, since the  $H_\alpha$  and  $H_\beta$  hydrogen spectral lines were recorded by the IGY patrol spectrograph at Fort Churchill whereas the helium line at 5876 Å was not.

The integral energy spectrum, derived from the pulse height distribution while the rocket was between 120 and 122 km, is shown in Figure 9. The left-hand curve has been corrected for the energy losses in the layer of aluminum in front of the detector, and the right-hand curve has been corrected for the energy losses in the aluminum and the atmosphere remaining above the rocket. The integral proton energy spectrum incident upon the atmosphere is seen to be approximately fitted by the function  $j(>E) = 2.5 \times 10^6 \exp(-E/30)$  protons/sec  $\text{cm}^2$  over the range 80 to 250 keV, where  $E$  is the proton energy in keV. The proton energy spectrum obtained in this manner is not dependent upon the model of the atmosphere used in extrapolating to above the atmosphere. The spectrum obtained from the proton flux versus altitude, however, is very strongly affected by the choice of the atmospheric model.

The atmospheric pressures at high altitudes given by various measurements and models differ by factors up to 4 or 5 to 1. Over the range 100 to 120 km, however, the ratios between these measurements and models are essentially

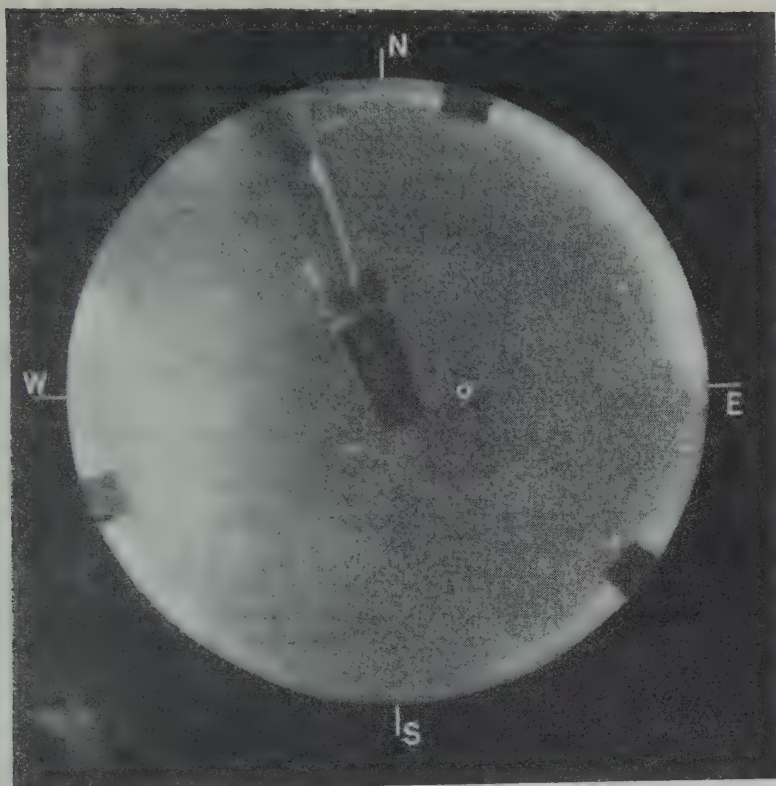


Fig. 6. All-sky photograph taken during the flight of rocket II 6.26F.

stant. By assuming that the atmospheric pressure was a factor of 1.5 less than that of the Rocket Panel [1952] atmosphere, it was found that the proton energy spectrum obtained from the altitude dependence of the proton detector output is similar to the spectrum obtained from the pulse height distribution. This assumed atmospheric pressure is about a factor of 1.3 greater than the day-time pressure measured at Fort Churchill by Horowitz, LaGow, and

Guiliani [1959]. For consistency, atmospheric pressures equal to those of the Rocket Panel model divided by 1.5 are used throughout this paper for pressures at altitudes above 95 km.

Figure 10 shows the proton spectrum derived from the pulse height distribution and from the altitude dependence during the rocket ascent and descent. The altitude and aspect curves shown in Figures 27 and 29 were used to obtain

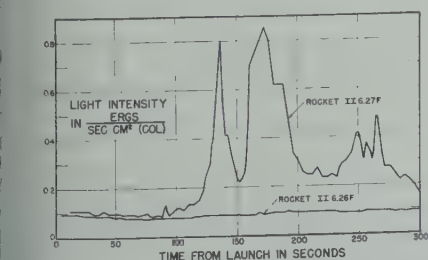


Fig. 7. Overhead light intensity measured by the ground-based photometer during the two rocket flights.

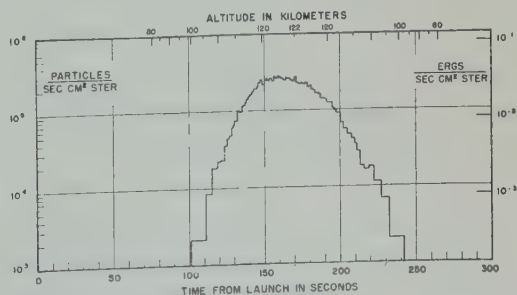


Fig. 8. Total flux of protons with energies greater than 45 keV measured by rocket II 6.26F.

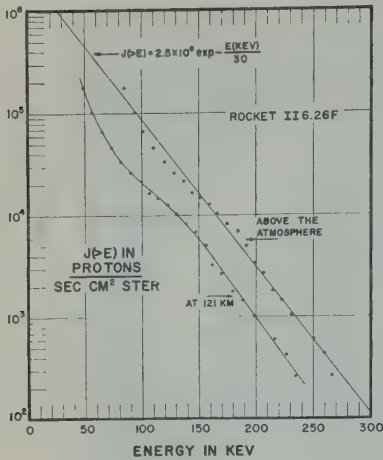


Fig. 9. Integral proton energy spectrum derived from the pulse height distribution obtained between 120 and 122 km by rocket II6.26F. The upper curve has been corrected for atmospheric absorption.

the atmospheric absorption of the protons. It was also assumed that the directional intensity of protons above the atmosphere was constant in time and isotropic between  $30^\circ$  and  $60^\circ$  with respect to the magnetic field. It can be seen from the curves in Figure 10 that the proton integral energy spectrum between 80 and 200 kev can also be represented by  $j(>E) = 2.5 (E/1000)^{-4.5}$  protons/sec  $\text{cm}^2$  ster, where  $E$  is in kev. This spectrum is similar in form to the one found for solar protons with energies between 30 and 300 Mev by *Anderson, Arnoldy, Hoffman, Peterson, and Winckler* [1959], but the constant for the auroral protons is lower by a factor of more than  $10^7$ . It is obvious, however, that the power-law representation of the solar proton spectrum cannot be extrapolated to energies of less than 1 Mev without predicting a flux of protons that is impossibly high. Interestingly, during the night of February 10-11, 1958, when solar protons were indirectly detected [*Reid and Leinbach*, 1959], the  $H_s$  intensity was measured by *Montalbetti* [1959] to be up to a factor of 10 higher than during the flight of rocket II6.26F.

The exponential form of the proton spectrum shown in Figure 9 predicts a total flux of about  $1.6 \times 10^7$  protons/sec  $\text{cm}^2$  when extrapolated to zero energy with the assumption of isotropy over the upper hemisphere. This extrapolation

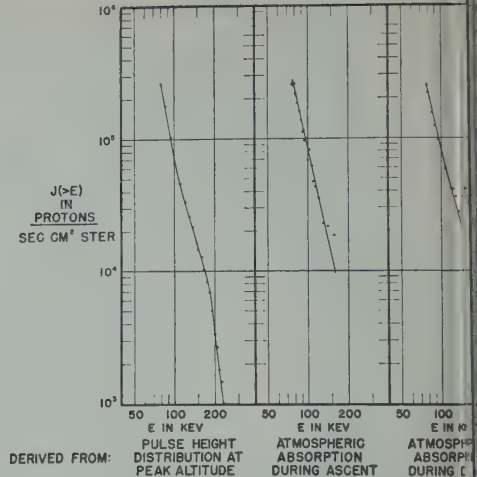


Fig. 10. Integral proton energy spectrum during the flight of rocket II6.26F. The points on the left-hand curve are the same as on the upper curve in Figure 9. The right-hand curves were derived from the number of protons with energies greater than 45 kev after atmospheric absorption.

would indicate that each proton produces about 4  $H_s$  quanta, which is in fair agreement with calculations made by *Chamberlain* [1954]. This result cannot be regarded as a direct measurement of the ratio of  $H_s$  quanta to incident protons since there is no reason for using any particular method for extrapolating total energy other than the fact that it gives a total proton flux.

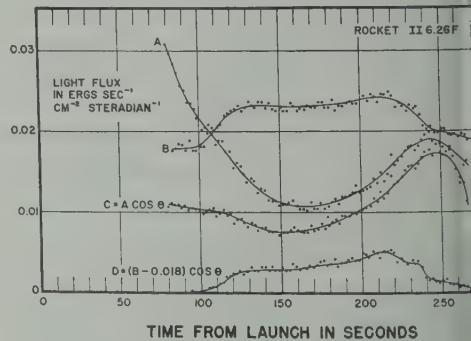


Fig. 11. Rocket photometer measurements during the flight of rocket II6.26F. Curves A and B represent the directional light intensity measured by the photometer when it was pointed close up and down, respectively. Curves C and D represent the measurements after correction. C gives the light intensity from auroral emission above and below the rocket.



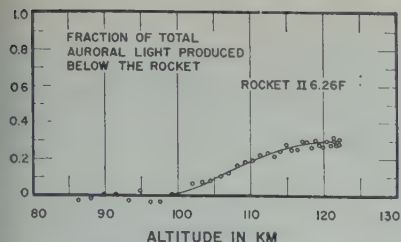


Fig. 12. Integral of the auroral luminosity versus altitude derived from curves *C* and *D* in Figure 11.

*Rocket measurements of auroral light.* The auroral light intensity measured by the rocket photometer during this flight is shown in Figure 11. Curves *A* and *B* are from measurements made when the photometer was directed closest up and down, respectively, during each spin cycle. Curves *C* and *D* are from the same measurements, but they are corrected for the angle of the photometer to the vertical. The downward measurements were also corrected for large background, presumably caused by auroral light reflected by the lower atmosphere and glow on the ground. The absolute intensity scale is uncertain by about a factor of 2, owing to lack of a reference level for the logarithmic response curve. Figure 12 is a graph of  $D/(D +$

*C*) from the measurements between 80 and 180 seconds shown in Figure 11 and therefore exhibits the fraction of auroral light produced below the rocket. This function is quite sensitive to the value taken for the background intensity. In particular, the slope in the vicinity of 120 km is subject to considerable error due to the low rocket velocity, which gave the background intensity more time to vary.

*Electrons.* The total energy flux versus time measured by the electron detector during this flight is shown in Figure 13. The dependence of the detector output upon the current in the focusing coil (see Fig. 23) indicates that most of this energy flux is due to electrons with energies of less than 40 kev.

The atmospheric absorption of the electron total energy flux can be used to obtain an estimate of the electron energy spectrum. If it is assumed that the light production efficiency of the atmosphere is constant between 100 and 120 km, the photometer measurements can also be used in finding the electron spectrum. It has been shown by Young [1956] and others that the average energy lost per electron by a beam of low-energy electrons in traversing an absorber is approximately equal to the energy  $E_0$  of an electron with an end-point range equal to the

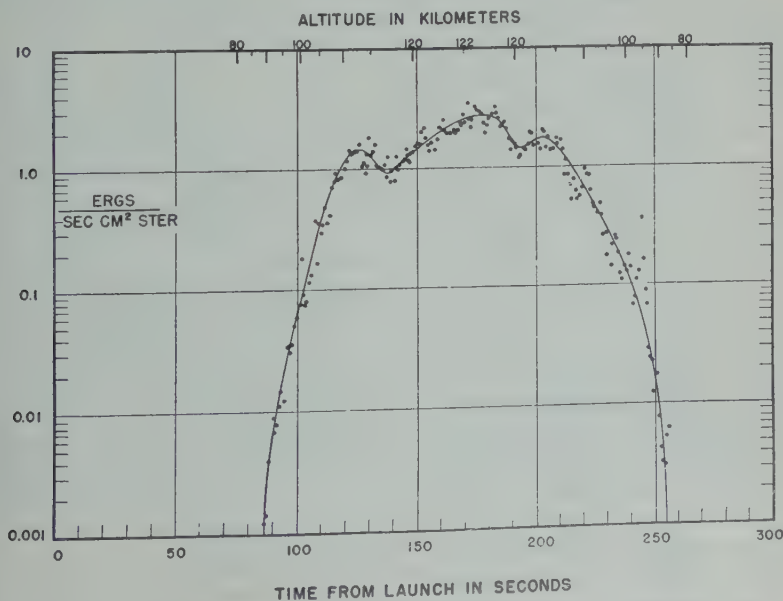


Fig. 13. Total energy flux versus time measured by the electron detector during the flight of rocket II 6.26F after atmospheric and instrumental absorption.

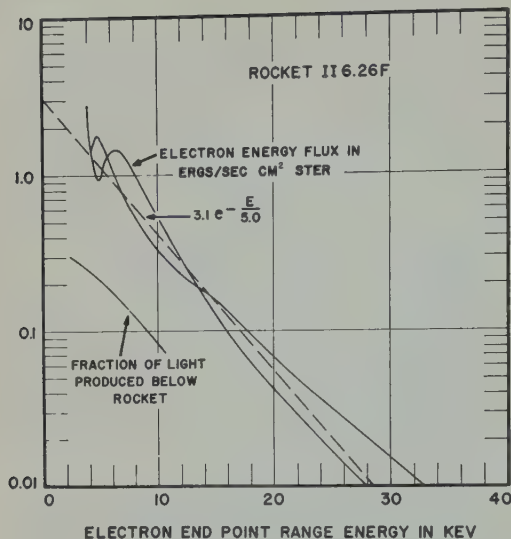


Fig. 14. The smooth curves in Figures 13 and 11 plotted against the electron energy which has a practical range equal to the atmospheric depth. The electron energy flux curve includes the additional absorption of the aluminum film covering the detector.

absorber thickness. In the present case of absorption by the high atmosphere in the presence of the relatively strong geomagnetic field, this approximation may not be good, but scattering probably controls the average electron trajectory more than the magnetic field does. If the integral number energy spectrum can be represented by a function of the form  $c \exp(-E/b)$ , where  $c$  and  $b$  are constants, the energy flux emerging from an absorber with an electron end-point range energy of  $E_0$  will be  $cb \exp(-E_0/b)$ . In Figure 14, the smooth curves of Figures 13 and 11 have been replotted against  $E_0$ .

The function  $3.1 \exp(-E_0/5.0 \text{ kev})$  ergs/sec  $\text{cm}^2$  ster fits these curves within the uncertainty caused by time variations. The integral number energy spectrum incident upon the atmosphere is therefore approximately  $2.5 \times 10^9 \exp(-E/5 \text{ kev})$  electrons/sec  $\text{cm}^2$  ster over the range 3 to 30 kev. The exact electron energy spectrum may have been considerably different from the result of this rather crude analysis, but it must have had the same general characteristics. The curve derived from photometer readings in Figure 14 extrapolates to a value of 0.5 at zero energy rather than to a value of

1.0. This discrepancy indicates either that the electron spectrum given above explains about one-half of the total energy flux or that the light production efficiency is higher in the very high atmosphere than in the 100- to 120 km region.

If it is assumed that the directional intensity of electrons was isotropic over the upper hemisphere, extrapolation of the above electron spectrum to zero energy indicates a total electron flux upon the atmosphere of about  $1.5 \times 10^{10}$  electrons/sec  $\text{cm}^2$  and a total energy due to electrons of about 20 ergs/sec  $\text{cm}^2$ .

The efficiency of the atmosphere between 100 and 120 km for converting energy lost by thermal electrons into visible light energy is found to be about 0.2 per cent. This result is based on the spectral sensitivity and absolute calibration of the ground photometer, the relative light produced below 120 km measured by the rocket photometer, and the assumption that the electrons were isotropic over the upper hemisphere. The probable error of this result is about a factor of 2. It is regarded as unlikely that the efficiency quoted above is in error by more than a factor of 4.

#### RESULTS FROM IGY ROCKET II 6.27F

*Behavior of the aurora and magnetic field.* IGY rocket number II 6.27F was fired at 0448 min 32 sec UT on February 25, 1958, in the presence of a bright auroral arc. Figure 15 is a photograph taken during the first 30 seconds of the rocket flight from a site 15 km west of the rocket launcher. The short streak, just above the horizon, left by the expended first stage, was produced by the ignition and burning of the second-stage rocket.

This aurora began as a bright homogeneous glow about 330 km to the south at 0455 UT and gradually extended without irregularities from horizon to horizon. At 0530 UT the aurora 'broke up' and the entire visible sky became brighter, and the aurora rapidly progressed northward in a regular fashion. At 0540 UT the bright northern edge of the aurora reached overhead at Fort Churchill. Coincident within  $\pm 3$  minutes of the auroral 'break up' magnetometer traces at all stations changed from very quiet to very disturbed. At Fort Churchill, the  $Y$  and  $Z$  components of the magnetic field varied rapidly

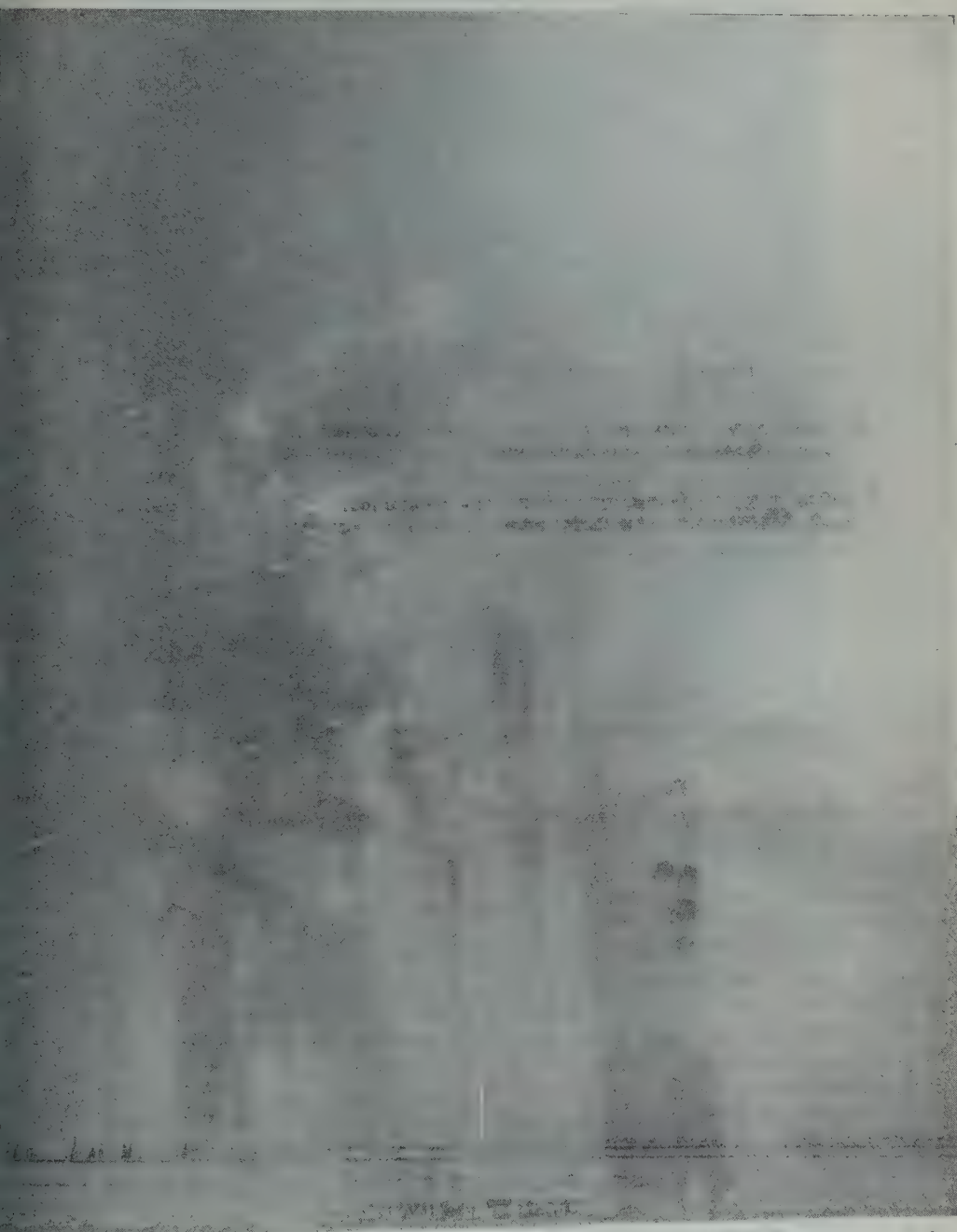


Fig. 15. Photograph taken during the first 30 seconds of rocket flight II627F from a site 15 km west of the rocket launcher. The traces produced during the burning periods of the first- and second-stage rockets are near the bottom and top of the photograph. The faint trace between these traces is produced by the expended first-stage rocket.



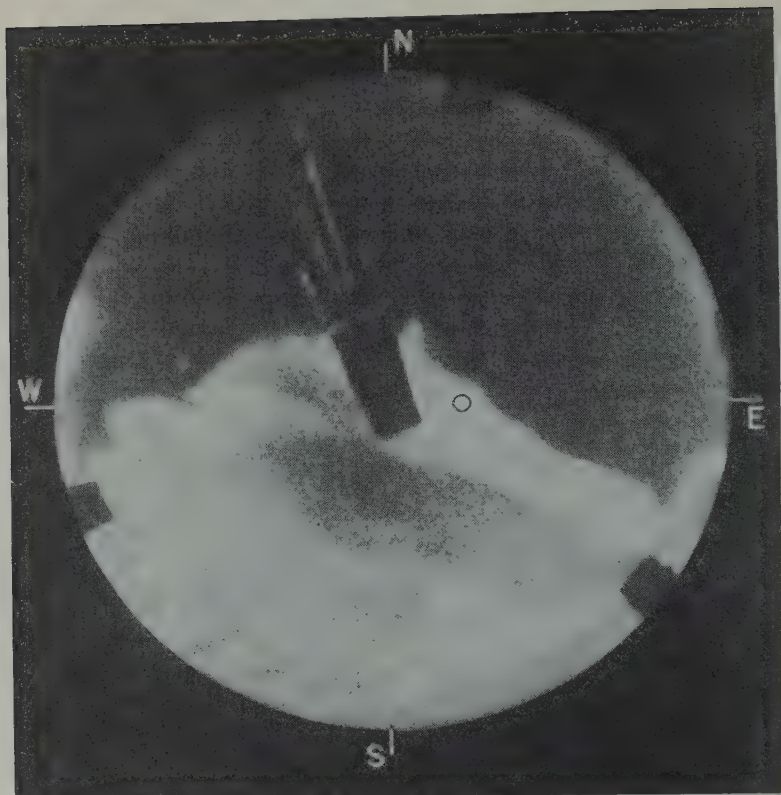


Fig. 16. All-sky-camera photographs taken at Fort Churchill 175 seconds after the launch of II6.27F. The position of the rocket at this time is indicated by the small circle.

over a range of about  $300 \gamma$ , and the  $X$  component disappeared, possibly off scale ( $> 1000 \gamma$ ). At O'Day, 130 km to the south of Fort Churchill, the single trace varied from 200 to  $400 \gamma$  from the quiet-period value. At Bird, 250 km south of Fort Churchill, the  $X$  component decreased about  $400 \gamma$ . Later the intense and rapidly varying auroral features filled most of the sky until 0700 UT, when the auroral brightness diminished. Fainter light emission persisted until dawn at 1100 UT.

Figure 16 is a photograph of this aurora taken at Fort Churchill 175 seconds after the launch of the rocket by a fast ( $f/0.71$ ) camera directed down toward a hemispherical mirror. The circle gives the position of the rocket at this time. Figure 17 is a similar photograph taken at O'Day (140 km southwest of the rocket) 140 seconds after launch. Figure 18 is a photograph taken at Belcher (100 km southwest of the rocket) between 142 and 147 seconds after launch; it is

one of many pairs of auroral-height-figures in photographs made at Belcher and O'Day [Ewen and Montalbetti, 1958]. The position of the rocket is again indicated by a circle. The height scale refers to the northern edge of the aurora and shows the lower boundary of the aurora at about 110 km high. For other azimuthal angles the scale must be modified to compensate for the different horizontal distances of the northern edge of the aurora.

The upper curve in Figure 7 shows the light intensity measured by the ground photometer during this flight. The light intensity measured by the rocket photometer shown in Figure 8 reveals that the auroral feature producing the peak at 135 seconds approached the rocket from the south but began to decrease in intensity just before reaching the rocket.

*Absence of protons.* In the number of protons above the discrimination point in the detector, this flight was similar to flight I.

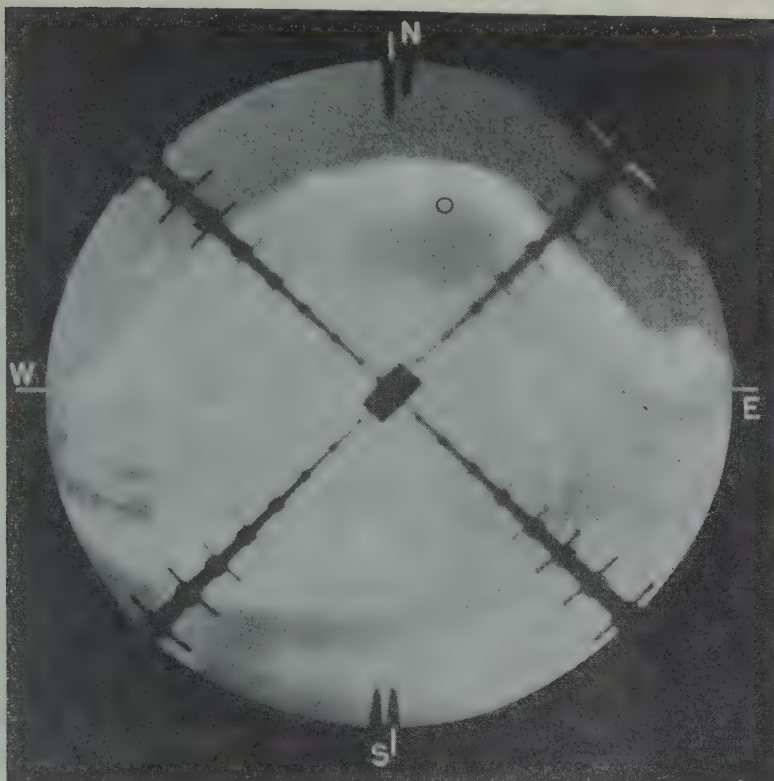


Fig. 17. All-sky-camera photograph taken at O'Day (140 km southwest of the rocket) 140 seconds after the launch of rocket II6.27F. The small circle indicates the position of the rocket.

Several factors combine, however, to indicate that in this flight most of these pulses were due to the light of the aurora. First, the quality of the layer of aluminum on the crystal was not as good as that on rocket II6.26F. Second, the discrimination point was lower, to the extent that a moderate number of photomultiplier noise pulses were counted even in total darkness. Third, the aurora was much brighter; and fourth, the flux of protons with greater than 100-kev energy was at least a factor of 10 less than during flight II6.26F. The pulse height distribution sharply decreased with increasing pulse height so that less than 1 out of every 50 pulses corresponded to energy losses in the crystal of greater than 50 kev. This indicates that less than  $4 \times 10^3$  protons/sec cm<sup>2</sup> ster with energies greater than 100 kev were incident on the atmosphere. No H<sub>β</sub> spectral line could be detected during this flight.

**Electrons.** The electron detector on this

flight measured total energy fluxes as high as 100 ergs/sec cm<sup>2</sup> ster. The energy flux versus time shown in Figure 20 in general follows the light intensity measured by the photometers but exhibits very rapid and irregular fluctuations. The output of the detector as a function of current in the focusing coil indicates that more than 90 per cent of this energy flux is due to electrons.

The energy spectrum of the electrons detected during this flight is very different from the electron energy spectrum measured on flight II6.26F. Instead of being distributed over a wide range of energies, the electrons in this case appear to have been nearly monoenergetic at about 6 kev. Since this energy is near the minimum energy detectable, the shape of the energy spectrum cannot be obtained with any precision. In the following paragraphs it will be shown, however, that less than 10 per cent of the total energy flux was due to electrons with greater

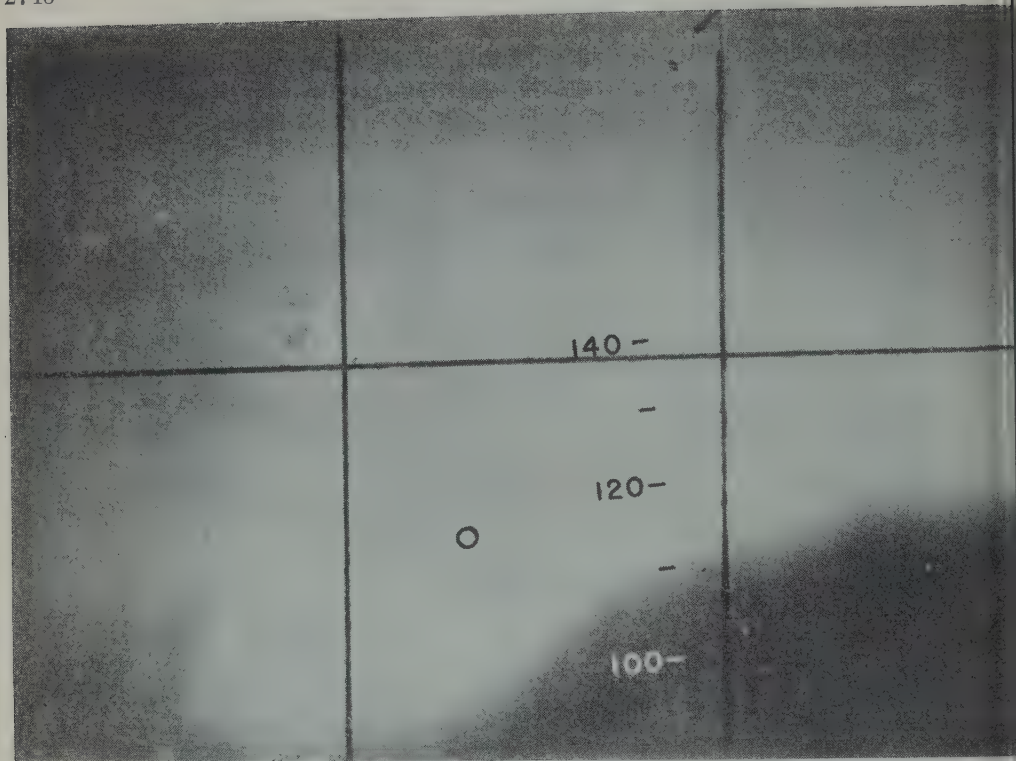


Fig. 18. Auroral height-finding photograph taken at Belcher between 142 and 147 seconds after launch of rocket II6.27F. The altitude scale in kilometers refers to the northern edge of the aurora.

than 10-kev energy and that less than 25 per cent of the auroral light was produced by electrons with less than 3-kev energy.

*Very-low-energy electrons.* The low performance of the rocket and the absorption by the aluminum on the crystal make it impossible to estimate the flux of electrons with less than 4-kev energy by means of the electron detector. The rocket photometer measurements can, how-

ever, be used to obtain some information about these very-low-energy electrons. Figure 21, obtained similarly to Figure 12, shows the fraction of auroral light produced below the rocket as a function of altitude. The scatter of the mean values is due to the rapid time variations, possibly to changes in the electron spectrum. These measurements show that the auroral luminosity sharply decreased below 112 km, in good agreement with the results obtained by auroral height-finding photographs as illustrated in Figure 18. The curve in Figure 21 is the integral of the auroral luminosity versus altitude; it predicts that relatively little light is emitted at altitudes greater than 130 km. The apparent light emission at higher altitudes in Figure 18 is probably from auroral light produced closer to the camera.

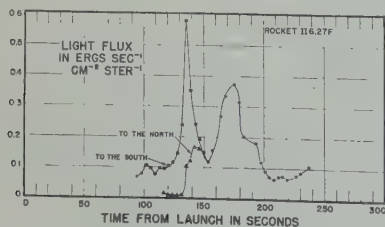
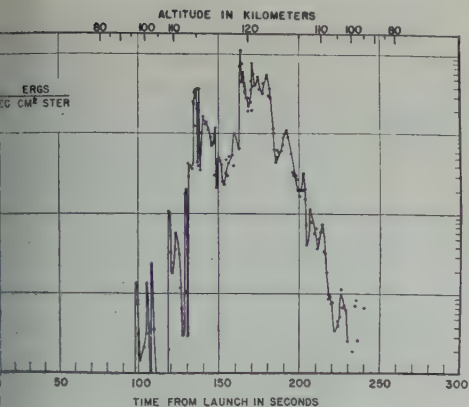


Fig. 19. Sum of the auroral light produced above and below rocket II6.27F measured by the rocket photometer. The light intensities looking slightly to the north and to the south were considerably different until 150 seconds after launch.

In the lower half of Figure 22 the data points shown in Figure 21 have been replotted with the altitudes converted to atmospheric density. The resulting curve is too close to linear to

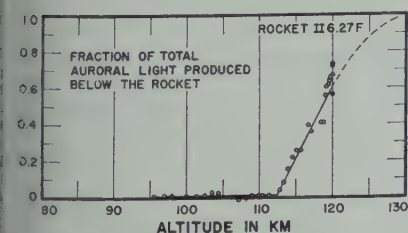




20. Total energy flux versus time during the flight of rocket II6.27F.

produced by electrons distributed in accordance to a simple power or exponential law but is very similar to the curve that might be expected from monoenergetic electrons. The initial energy of the electrons, corresponding to the range of  $8 \times 10^{-5}$  g/cm<sup>2</sup> given by this curve, could have been between 4 and 7 kev, depending on the electron angular distribution assumed and the model of the atmosphere used. Approximately 50 per cent of the auroral light is produced by these electrons in the last half of their range (40–80  $\mu$ g/cm<sup>2</sup>). Therefore, these electrons probably produced at least 25 per cent of the auroral light in the first half of their range (0–40  $\mu$ g/cm<sup>2</sup>), leaving less than 25 per cent of the auroral light to be produced by electrons with less than 3-kev energy.

**Total electron flux.** To obtain the dependence of the total energy flux upon atmospheric depth, the measurements shown in Figure 14 must be corrected for the large time variations.



21. Integral of the auroral luminosity versus altitude derived from the rocket photometer measurements between 160 and 240 seconds after launch of rocket II6.27F.

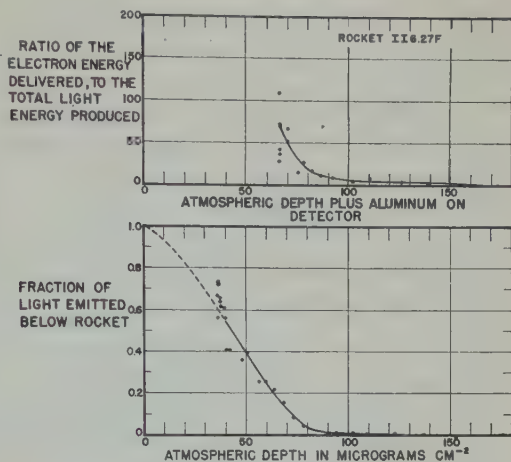


Fig. 22. The measurements shown in Figure 21 are plotted versus the atmospheric depth in the lower graph. The upper graph shows the ratio of the particle energy delivered to the crystal in the electron detector to the total visible light energy produced by the atmosphere as measured by the rocket photometer.

The rocket photometer measurements shown in Figure 19 give a crude indication of the time variations, and they were used to obtain the ratio of the electron energy delivered to the electron detector scintillator to the total light being produced along the line of force. These ratios, plotted against the total thickness of aluminum on the detector and atmosphere above, are shown in the upper half of Figure 22. Comparison with the lower half of the figure indicates that only about 20 per cent of the energy flux incident upon the atmosphere reached the scintillator when the rocket was at peak altitude. Assuming the directional intensity of electrons at the rocket to be constant over the upper hemisphere, the electron flux incident upon the atmosphere between 165 and 180 seconds after launch was approximately  $2 \times 10^{11}$  electrons/sec cm<sup>2</sup> if the energy per electron was 6 kev. This corresponds to a total energy flux of about 2000 ergs/sec cm<sup>2</sup>. The possible error in assuming isotropy over the upper hemisphere is probably not serious, regardless of the initial angular distribution, because of atmospheric scattering and the fact that the detector averaged the flux from over a wide range of angles.

*Absence of high-energy electrons.* The rela-

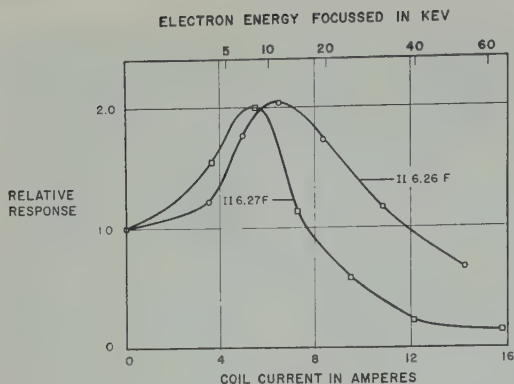


Fig. 23. Relative output of the electron detector as a function of current in the focusing coil during the two rocket flights.

tive absence of electrons with energies greater than 7 kev can be deduced from Figure 22, which shows that only about 1 per cent of the initial energy flux remained after absorption by  $100 \mu\text{g}/\text{cm}^2$  of material.

Comparison of the relative output of the electron detector versus coil current shown in Figure 23 with that of rocket II 6.26F also shows the relative absence of higher-energy electrons. The points on these curves have an accuracy of only about  $\pm 15$  per cent, owing to the wide dynamic range covered by the analog output and to time variations during changes in coil current. The relative output of the detector versus coil current for isotropic monoenergetic electrons can be obtained in the laboratory by measuring the geometrical factor at different currents with a parallel pencil of monoenergetic electrons directed at the detector from different angles and positions. For practical reasons, the accuracy obtained to date with this procedure is rather poor, but the resulting function is very similar to the curve obtained during the flight of rocket II 6.27F. Without an accurate knowledge of this function, the conversion of the curves in Figure 23 into electron energy spectra is impossible. A qualitative picture of the differential number energy spectra implied by these curves can be obtained by noting that the relative sensitivity of the detector for focused electrons increases approximately as the fourth power of the coil current. This can be seen from the following facts: The energy loss per electron

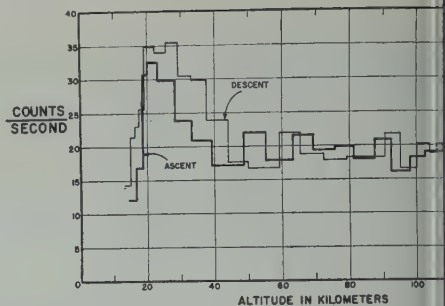


Fig. 24. Geiger tube counting rate versus altitude during rocket flight II 6.26F.

in the scintillator increases as the first of the electron energy; the differential magnetic rigidities focused is proportional to coil current; the electron energy focused is approximately proportional to the square of the magnetic rigidity focused; and the rigidity focused is directly proportional to coil current.

Further evidence for the relative abundance of higher-energy electrons during this flight is the fact that the Geiger tube counted only low-energy rays. In the case of rocket II 6.26F, the spectrum of  $2.5 \times 10^6 \exp(-E/5)$  predicted by the measured efficiency of the detector predicts a contribution to the counting rate of about 3 counts/sec. Figure 24 shows the counting rate of the Geiger tube in rocket flight II 6.26F versus altitude. The average rate above 40 km of  $19.0 \pm 0.3$  counts/sec is to be compared with the average of 16.6 counts/sec obtained during other rocket flights at Fort Churchill. The absence of a significant increase above 40 km, however, suggests that the difference

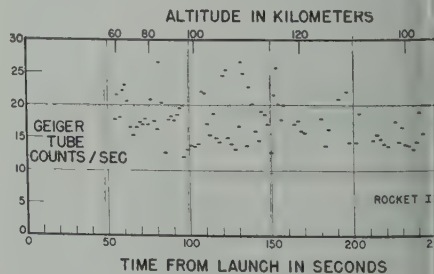


Fig. 25. Geiger tube counting rate versus time from launch during the flight of rocket II 6.27F. The line to the right gives the probable error of individual measurements.

counts/sec may have been caused by an increase in the cosmic-ray intensity.

Figure 25 shows the counting rate of the Geiger counter in rocket II6.27F versus time. The short bars represent the average counting rate for 32 counts received, giving each bar a probable error of 3 counts/sec from the average rate of 17.0 counts/sec. The lack of any statistically significant variation from the expected cosmic-ray counting rate implies that the intensity was less than during the flight of rocket II6.26F even though the total energy flux was much as 50 times greater.

### DISCUSSION

*Electrons incident upon the atmosphere.* Examination of the results presented in the preceding sections indicates that the luminosity of the auroras investigated was primarily due to the incidence of energetic electrons upon the atmosphere. In each, the absolute intensity and the distribution of the auroral emissions could easily have been produced by the electrons detected by the rocket-borne instruments. The acceleration in the light-emitting regions, as in an atmospheric discharge, was therefore either absent or not important in the production of the aurora.

*Intensity of hydrogen spectrum.* Chamberlain [1958] estimates that a 90-keV proton produces  $1.5 \times 10^3$  quanta in its total path length in the atmosphere, therefore converting about  $10^{-3}$  of the proton energy into  $H_\alpha$  quanta. The conversion efficiency would be higher for lower-energy protons. This efficiency is to be compared with the measured efficiency of  $2 \times 10^{-3}$  for the conversion of electron energy into all forms of light by the atmosphere between 100 and 120 km. It would appear that the  $H_\alpha$  auroral line, in an aurora produced by protons, should be almost as intense as the O I 5577 Å. Since the relative  $H_\alpha$  intensity has never been reported to be this high, most auroral emissions in the 100-to 120-km region are probably due to electrons. Omholt [1959] also concludes, on the basis of the observed ratios of  $H_\alpha$  line intensity to the  $N_2^+$  band intensity of 100 Å, that protons regularly supply only a small part of the energy necessary to produce the aurora.

*Effective temperatures.* The electron and proton energy spectra measured during the flight of rocket II6.26F can be represented within the accuracy of the measurements by Maxwell-Boltzmann energy distributions. The effective temperatures obtained are  $(5.4 \pm 1) \times 10^7$  °K for the electrons and  $(3.2 \pm 0.4) \times 10^8$  °K for the protons.

Many theories can be conjectured which give a qualitative explanation of these energy spectra [Chamberlain, 1958]. The choice between them must await further experimental results. Measurements of the energy spectra of the solar particles while in transit to the earth would be particularly valuable in this regard.

The electron energy spectrum measured during the flight of rocket II6.27F cannot be represented by a Maxwell-Boltzmann energy distribution. As was mentioned previously, the electron energy spectrum during this flight cannot be determined accurately from the available data. It was shown, however, that the spectrum decreased precipitously toward higher energies in the region of 10 keV but did not increase rapidly toward lower energies in the region below 4 keV. Quantitatively, this spectral form could be represented by a Maxwell-Boltzmann energy distribution with a temperature of about  $6 \times 10^7$  °K except for the fact that far too few high-energy electrons were present. In common with the Maxwell-Boltzmann distribution, almost all particle energy spectra produced by acceleration mechanisms involving statistical processes predict a moderate number of particles with energies considerably higher than the average particle energy. Acceleration processes of a statistical nature therefore could not have been important in the history of the particles producing the aurora penetrated by rocket II6.27F. Acceleration processes involving electric fields might, however, produce an electron energy spectrum with a sharp high-energy cutoff.

*Magnetic field instabilities.* The rapid variations and distorted appearance of the aurora penetrated by rocket II6.27F are strong evidence for instabilities in the magnetic field somewhere along these lines of force.

The number of electrons incident upon the atmosphere during these flights was not sufficiently high to produce these instabilities. In



order to reach the atmosphere, however, auroral particles must be aligned to within  $\pm 2^\circ$  of the line of force in the vicinity of the magnetic equator. The particles producing auroral light are therefore probably only a small fraction of the total number of particles involved.

It is of interest to see whether the particles whose trajectories make angles greater than  $\pm 2^\circ$  to the line of force at the equator could have produced magnetic instabilities by assuming that their directional intensity was equal to the average directional intensity reaching the atmosphere. Whether the particle energy density is less or greater than the magnetic field energy density is often used as a criterion of whether the field is stable or unstable. For rocket II6.26F, this criterion would not be quite met even if the measured directional particle intensity is assumed to have extended over the full  $4\pi$  steradians. For rocket II6.27F, however, an extension of the measured intensity over a solid angle of only about 0.3 steradian would have been sufficient to meet this criterion for instability. This result is consistent with the relative stability of the auroral features during the flight of rocket II6.26F and the indications of instability during the flight of rocket II6.27F.

*Break-up of quiescent arcs.* The particles with trajectories that do not reach the atmosphere are reflected back toward the equator by the converging magnetic field. If the injection mechanism did not also act as an efficient loss mechanism, the particles would be trapped and accumulate until the magnetic field could no longer contain them. The magnetic field would therefore become unstable. The ensuing rapid variations of the magnetic field would tend to perturb the trapped-particle trajectories into directions which result in the loss of particles by absorption in the atmosphere.

The build-up of trapped particles may correspond to the initial quiet arc phase of an auroral display and the subsequent break-up to the onset of magnetic instabilities. The rapid variations of the magnetic field would also tend to raise some of the particles to higher energies by Fermi-type accelerations. The particles detected by rocket II6.27F may have been freshly injected particles whose energy spectrum had been relatively unmodified by post-injection accelerations. The particles detected by rocket

II6.26F may have been older particles whose energy spectrum had been extended to higher energies while they were trapped in the magnetic field.

*Comparison with results from IGY NN3.03F.* Concurrent with the rocket investigation of auroras reported in this paper, auroral experiments were performed by Meredith, and Berg of the U. S. Naval Research Laboratory (now with the National Aeronautics and Space Administration). The results obtained by these investigators during the flight of IGY NN3.03F [Meredith, Davis, Heppner, and Berg, 1958; Davis, Berg and Meredith, 1960] are remarkably similar to those obtained during the flight of rocket II6.26F. In particular, the electron energy flux and the intensity of particles with energies greater than 70 kev were within a factor of 3 of the corresponding measurements made during the flight of rocket II6.26F.

*Auroral-zone X rays.* The absence of X-ray Geiger tube counting rates during these flights is in sharp contrast with the results obtained during previous auroral zone rocket flights [Meredith, Gottlieb, and Van Allen, 1955; Van Allen, 1957]. The high Geiger tube counting rates during the previous flights, however, can easily be explained by fluxes of electrons comparable in number to those during the present flights if it is assumed that the effective temperature of the electrons was about a factor of 2 higher than during the flight of rocket II6.26F.

*Comparison with trapped radiation.* It has been suggested that auroral particles are closely related to particles trapped in the outer radiation zone [Van Allen, McIlwain, and Wibberley, 1959]. Since, to a certain extent, the phenomena occupy the same region of space, some relationship is to be expected, but the nature remains obscure.

The total intensity of trapped electrons at the center of the outer zone [Van Allen, Frank, 1959; Arnoldy, Hoffman, and Lerger, 1960] is similar to the total flux of electrons producing visible auroras. The energy spectrum of the trapped electrons is a moderately smooth function of time and spatial position. Most of the observations, however, can be explained by the spectral form  $c \exp(-E/70 \text{ kev})$  measured at low altitudes near the inner edge of the

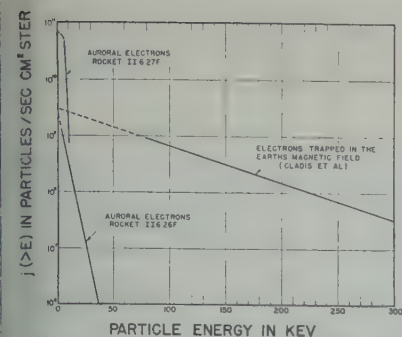


Fig. 26. Comparison of the integral electron spectra obtained by the two auroral rocket flights with the spectrum that may be present in the outer zone of trapped radiation.

by Cladis, Chase, and Imhof [1960; Walt, Cladis, Imhof, and Knecht, 1960]. It is interesting to note that the inner-zone electrons have a similar spectral form [Holly and Johnson, 1960]. These measurements of the trapped electron energy spectrum can also be represented by a Maxwell-Boltzmann distribution, but with a temperature approximately an order of magnitude higher than the auroral electrons. The trapped-electron energy spectrum is compared with the auroral energy spectra in Figure 26. The version of the processes suggested previously to explain the behavior of auroras may be responsible for the trapped electrons, but many more detailed experiments must be performed in order to establish the origin of these particles.

#### FUTURE EXPERIMENTS

Auroral features take a great number of forms, many of which cannot be explained by the modifications of the particle fluxes ob-

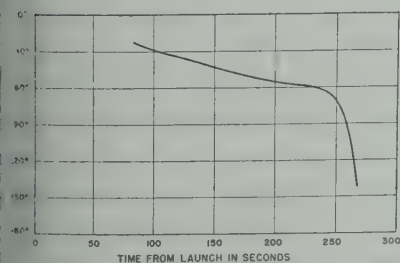


Fig. 27. The angle between the particle detector and the negative of the magnetic field vector during the flight of rocket II 6.26F.

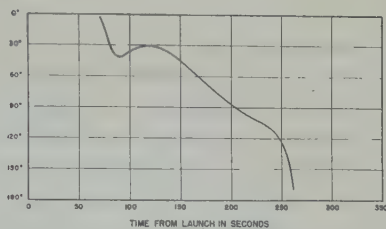


Fig. 28. The angle between the particle detectors and the negative of the magnetic field vector during the flight of rocket II 6.27F.

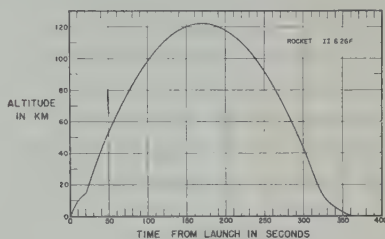


Fig. 29. Altitude of rocket II 6.26F versus time.

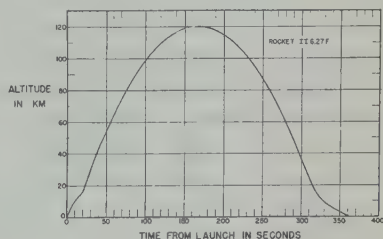


Fig. 30. Altitude of rocket II 6.27F versus time.

served producing the auroras described in this paper. A properly instrumented earth satellite in a high-inclination orbit would be invaluable in determining the nature of the particles producing the different auroral features.

The detectors in future rockets and satellites should be capable of measuring the absolute energy spectra and angular distribution of electrons and protons over the energy range of at least 1 keV to 1 MeV. Simultaneous measurements with such instruments of the particles in transit from the sun, of the particles in the regions where the auroral particles are injected (and/or accelerated), and of the particles producing a visible aurora should be capable of revealing the true nature of this bewildering phenomenon.

## CONCLUSION

The light emitted by the atmosphere below 120 km during two different auroras was found to be produced by low-energy electrons. Protons were detected in one, but the intensity was too low to contribute appreciably to the light emitted below 120 km.

The electron and proton energy spectra derived from measurements made during the flight of rocket II6.26F can be represented by Maxwell-Boltzmann distributions with temperatures of  $5.4 \times 10^4$  and  $3.2 \times 10^4$  degrees Kelvin, respectively. The energy spectrum of the electrons producing the aurora penetrated by rocket II6.27F, however, did not have a high-energy tail. The acceleration mechanism producing these nearly monoenergetic electrons therefore presumably did not involve a statistical type process.

During both rocket flights there was a striking absence of electrons with greater than 50-kev energy. The fact that the rocket-borne Geiger tube counting rates were less than a factor of 1.2 above the expected cosmic-ray rate is in great contrast to the very high counting rates anticipated on the basis of previous auroral-zone rocket flights. The total electron flux in these auroras was similar to the electron flux in the outer zone of trapped radiation, but the individual electron energies of the trapped electrons are typically an order of magnitude higher.

**Acknowledgments.** I wish to thank Professor James A. Van Allen for making this investigation possible and for his suggestions and guidance in the construction of the instruments and the analysis of the results.

The assistance of Mr. D. C. Enemark in the design and preparation of the many electronic circuits is gratefully acknowledged.

Many important optical observations of the auroral features were made by the Defence Research Northern Laboratory of the Canadian Department of National Defence. In particular I wish to thank Dr. R. Montalbetti, Mr. D. McEwen, and Mr. H. Lutz for obtaining the  $H_\beta$  spectral intensity and many photographs of the auroras penetrated by the rockets.

I wish to express appreciation to the many U. S. Army and Navy personnel who assisted in the rocket-launching operations.

I am grateful to Mr. H. Hills and Mrs. A. Hudmon for the reduction and analysis of the telemetry records.

The investigation was supported by the National Academy of Sciences and the National Science

Foundation through their US/IGY Project and by the joint program of the Office of Research and the Atomic Energy Commission.

## REFERENCES

- Anderson, K. A., R. Arnoldy, R. Hoffman, Peterson, and J. R. Winckler, Observation of low-energy solar cosmic rays from the flare of August 1958, *J. Geophys. Research*, **64**, 1147, 1959.
- Arnoldy, R. L., R. A. Hoffman, and J. R. Winckler, Observation of the Van Allen regions of August and September 1959, I, *J. Geophys. Research*, **65**, 1361-1376, 1960.
- Chamberlain, J. W., The excitation of hydrogen aurora, *Astrophys. J.*, **120** 360-366, 1954.
- Chamberlain, J. W., Theories of the aurora, *Advances in Geophys.*, **4**, 109-215, Academic Press, New York, 1958.
- Cladis, J. B., L. F. Chase, and W. L. Melrose, Energy spectra and spatial distributions of electrons trapped in the earth's magnetic field, *Am. Phys. Soc.*, [2] **5**(1), 46, 1960.
- Davis, L. R., O. E. Berg, and L. H. Meredith, Direct measurements of particle fluxes in auroral zones, *Proc. Cospar Space Sci. Symp.*, North Holland Publishing Co., Amsterdam, 1960.
- Havens, R. J., R. T. Koll, and H. E. LaGorce, pressure, density, and temperature of the atmosphere to 160 kilometers, *J. Geophys. Research*, **57**, 67, 1952.
- Holly, F. E., and R. G. Johnson, Measurements of radiation in the lower Van Allen belt, *J. Geophys. Research*, **65**, 771-772, 1960.
- Horowitz, H. E. LaGorce, and J. Guiliani, First measurements of auroral-zone atmospheric structure from 100 to 188 km, *J. Geophys. Research*, **64**, 2287, 1959.
- Lockwood, J. A., Decrease of cosmic-ray intensity on February 11, 1958, *J. Geophys. Research*, **64**, 27-37, 1960.
- McEwen, D. J., and R. Montalbetti, Particle measurements on aurora over Churchill, Canada, *Can. J. Phys.*, **36**, 1593-1600, 1958.
- Meredith, L. H., L. R. Davis, J. P. Heppner, O. E. Berg, Rocket auroral investigation, *Rocket Rept. Ser. 1*, 169-178, National Academy of Sciences, National Research Council, Washington, D. C., 1958.
- Meredith, L. H., M. B. Gottlieb, and J. Van Allen, Direct detection of soft radiation at 50 kilometers in the auroral zone, *Phys. Rev.*, **131**, 201-205, 1955.
- Montalbetti, R., Photoelectric measurements of hydrogen emissions in aurora and airglow, *Atmospheric and Terrest. Phys.*, **14**, 200, 1960.
- Ohmolt, A., Studies on the excitation of hydrogen, I, The hydrogen lines, *Geofys. Papperer*, **20**(11), 1-40, 1959.
- Reid, G. C., and H. Leinbach, Low-energy electron ray events associated with solar flares, *J. Geophys. Research*, **64**, 1801-1805, 1959.



- et Panel, Pressures, densities, and temperatures in the upper atmosphere, *Phys. Rev.*, **55**, 1027-1032, 1952.
- ner, C., *The Polar Aurora*, Oxford University Press, London, 1955.
- Allen, J. A., Direct detection of auroral radiation with rocket equipment, *Proc. Natl. Acad. Sci. U. S. A.*, **43**, 57-92, 1957.
- Allen, J. A., and L. A. Frank, Radiation measurements to 658,300 kilometers with Pioneer IV, *Nature*, **84**, 219-224, 1959.
- Allen, J. A., C. E. McIlwain, and G. H. Ludwig, Radiation observations with satellite 1958e, *Geophys. Research*, **64**, 271-286, 1959.
- Walt, M., L. F. Chase, J., J. B. Cladis, W. L. Imhof, and D. J. Knecht, Energy spectra and spatial distributions of electrons trapped in the earth's magnetic field, *Proc. Cospar Space Sci. Symposium*, North Holland Publishing Co., Amsterdam, 1960.
- Winckler, J. R., L. Peterson, R. Hoffman, and R. Arnoldy, Auroral X-rays, cosmic rays, and related phenomena during the storm of February 10-11, 1958, *J. Geophys. Research*, **64**, 597-610, 1959.
- Young, J. R., Penetration of electrons and ions in aluminum, *J. Appl. Phys.*, **27**, 1-4, 1956.

(Manuscript received June 10, 1960.)



# Audio-Frequency Electromagnetic Radiation in the Auroral Zone

GEORG GUSTAFSSON AND ALV EGELAND

*Kiruna Geophysical Observatory  
Kiruna, Sweden*

AND

JULES AARONS

*Propagation Sciences Laboratory  
Electronics Research Directorate  
Air Force Cambridge Research Center  
Bedford, Massachusetts*

**Abstract.** During three one-month periods, continuous spectrograms of the electromagnetic energy in the spectral region between 10 cps and 10 kc/sec were recorded in Kiruna, Sweden. The records were examined from the viewpoint of background energy. Throughout the entire frequency band studied, there is a low daytime signal level and a nighttime maximum. The ratio of the maximum to minimum amplitude varies as a function of frequency (the higher maximum-to-minimum ratio occurs at the lower frequency range of 20 to 200 cps) and with the season of the year. It is concluded that the daytime ionosphere absorbs the energy throughout the entire spectrum studied. It was found that, although strong deviations of the signal level from the normal were often associated with geomagnetic disturbances, there was a general lack of correlation between magnetic index and low-frequency noise, except in the 0- to 45-cps frequency range.

The origin of the background signals is probably twofold, atmospherics from great distances as well as magnetic and exospheric fluctuations contributing to the lower band.

On eleven occasions, electromagnetic radiation associated with micropulsations of the earth's magnetic field was detected. Two frequency bands were identified: one centered at 750 cps, which is the gyro frequency for protons at an altitude of 100 km above Kiruna; and the other ranging between 1.8 and 4.5 kc/s, which has been identified as hiss. On all but one of the occasions when emissions were detected, the 750-cps signals were quite stable in frequency. During four of the longest periods when radiation was recorded, the low-frequency emissions were received between two phases of a magnetic storm; micropulsations were simultaneously evident on the Kiruna magnetograms.

## 1. INTRODUCTION

The existence of naturally occurring electromagnetic radiation in the audio-frequency range was first discussed by *Menzel and Salisbury* [1947], referring to observations in the range 100 to 400 c/s. Since then, the audio-frequency emissions have been investigated experimentally, and the origin of the low-frequency noise has been discussed, by several authors [*Willis*, 1948; *Woods and Henissart*, 1953; *Holtzer and Deal*, 1954; *Aarons*, 1954, 1956; *Goldberg*, 1956; *Wilford and Maple*, 1957; *Gallet*, 1959; *Ellis*, 1959]. In all investigations, except those of *Maple* [1957] and *Egeland* [1959], were carried out over latitudes, and only the frequency range 100 to 400 c/s was studied.

During three periods of approximately 1 month each (October 12 to November 11, 1958; December 12, 1958, to January 11, 1959; and April 15 to May 14, 1959), continuous recordings were taken at the Kiruna Geophysical Observatory, Sweden (geographical coordinates, 67.8N, 20.4E; geomagnetic coordinates, 65.3N, 115.5E), of the electromagnetic energy in the audio-frequency range.

The amplitudes of audio-frequency signals were recorded as a function of frequency and time. The aim of the program was to investigate the diurnal and seasonal variation as a function of time and frequency, to investigate the possibility of emission and absorption of low-frequency energy in connection with strong



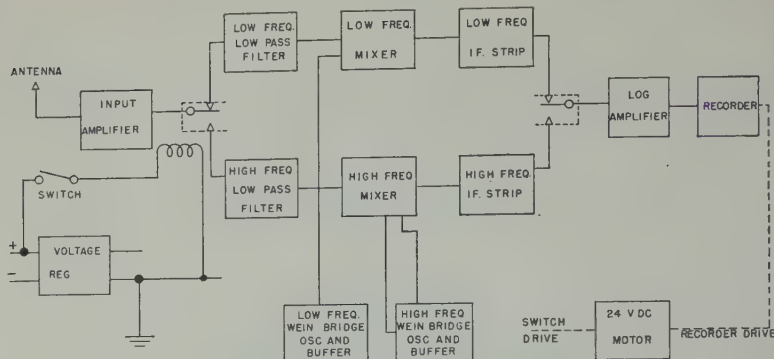


Fig. 1. Diagram of narrow-band sweep-frequency analyzer used in the investigation.

magnetic and auroral activity (e.g., emission of electromagnetic energy from protons at the gyro frequency in the upper atmosphere), and to compare the geophysical phenomena and solar activity.

## 2. INSTRUMENTATION AND OBSERVATION SITE

The equipment used is a narrow-band sweep-frequency analyzer, a block diagram of which is shown in Figure 1. It is completely transistorized, powered by a 24-volt battery. The power consumption is less than 10 watts. Each sweep is divided into two frequency ranges: 10 to 1000 c/s, and 0.5 to 10 kc/s, with bandwidths of 1 c/s and 40 c/s, respectively. The output

from the log amplifier is recorded on an Angus recorder, driven by a motor which also varies the frequency setting potentiometer.

The pickup is a 30-foot vertical antenna. The equipment is housed 6 km from the Observatory and approximately 2 km from the nearest town.

## 3. SCALING TECHNIQUE

Amplitude and frequency variations have been investigated. In analysis, each sweep of a duration of 32.8 minutes, is divided into the following frequency intervals: 20 to 200 c/s, 200 to 700 c/s, and 3 to 10 kc/s. An example of a registration with the recording milliammeter is shown in Figure 2.

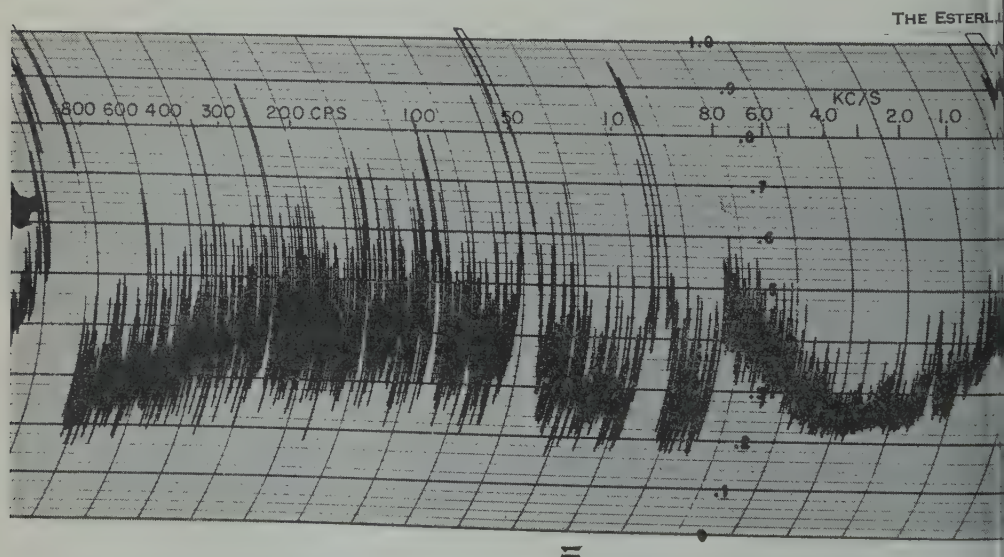


Fig. 2. Sample recording of sweeps 10 c/s to 1 kc/s; 0.5 kc/s to 10 kc/s, Dec. 2, 1958, 0030.

be seen in Figure 2. The maximum amplitude and its frequency have been scaled within three intervals. According to Wilcox and Le [1957], the maximum amplitude seems to represent the data quite well. The amplitude measured in equivalent voltage output from the antenna. Peaks at the harmonics of the carrier line up to 750 c/s have been omitted from the data, but, owing to the narrow filter bandwidth, little information is lost. Calibration has been done almost every day with a General Radio microvolter at the input of the amplifier.

#### 4. DIURNAL VARIATION IN AMPLITUDE

The field strength of the low-frequency emission has been reduced. The signal characteristics show a high degree of variability during a 24-hour interval, but usually in a relatively systematic way.

The diurnal variation in the frequency range 20 to 200 c/s. The diurnal variation for three months' records is shown in Figure 3. The variation in signal strength exhibits a well-defined maximum around local midnight and a minimum around sunrise and sunset. The short vertical lines in Figure 3 indicate the average time of sunrise and sunset in the horizon. The minimum intensity occurred around noon in the winter months but much earlier in the spring (April-May). In the records from the December-January period, a small secondary maximum between 1200 and 1500 MET can be seen. A surprising feature is the very high signal intensity found between 1200 and 1800 MET in the last period.

The ratio of maximum to minimum signal strength for the three periods is 2:1, 1.5:1, and 1.1:1 respectively, as can be seen in Figure 3. Another important feature of the curve is the regular morning decrease at sunrise.

The diurnal variation for the range 3 to 10 kc/s. Figure 4 shows the diurnal variation in the higher-frequency range for the same periods of investigation. The histograms are obtained by averaging the observations for each period of 30 minutes. For the first two periods the curves in Figure 4 show approximately the same diurnal variation in occurrence as the lowest-frequency interval. But the pattern is quite different during the April-May period, where no distinct

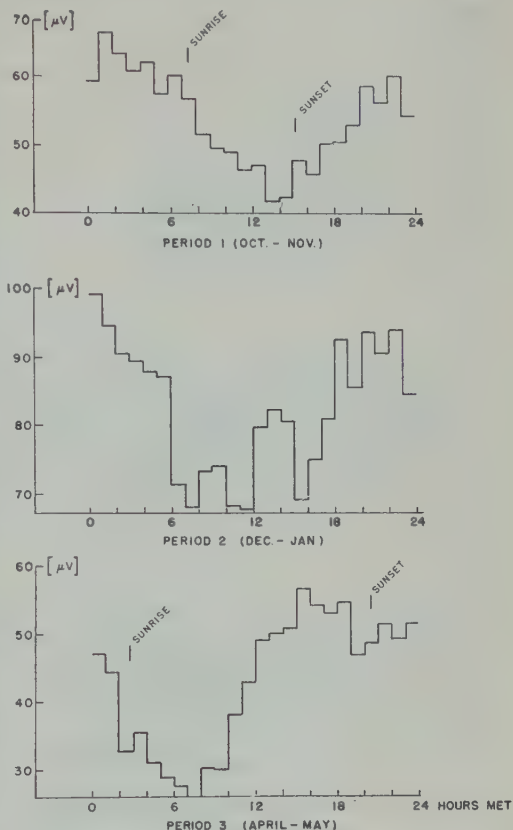


Fig. 3. Diurnal variation in the frequency range 20 to 200 c/s.

diurnal variation is apparent. The character of the record, however, usually changes from hour to hour in a more irregular fashion. For the first two periods, the maximum intensity is found in the nighttime with a definite minimum between 1000 and 1600 MET. In this frequency range a decrease shortly after sunrise is observed. The ratio of maximum to minimum signal level for the three curves in Figure 4 is 1.6:1, 1.8:1, and 1.1:1.

The diurnal variation of the signal intensity on October 24, 1958, in the range 3 to 10 kc/s is also illustrated in Figure 5. The time between recordings is approximately 4 hours. The increase of the whole level between 4.5 and 10 kc/s during the night is obvious, whereas in the range 3 to 4.5 kc/s the level is constant during the same time.

*Diurnal variation for the frequency range*

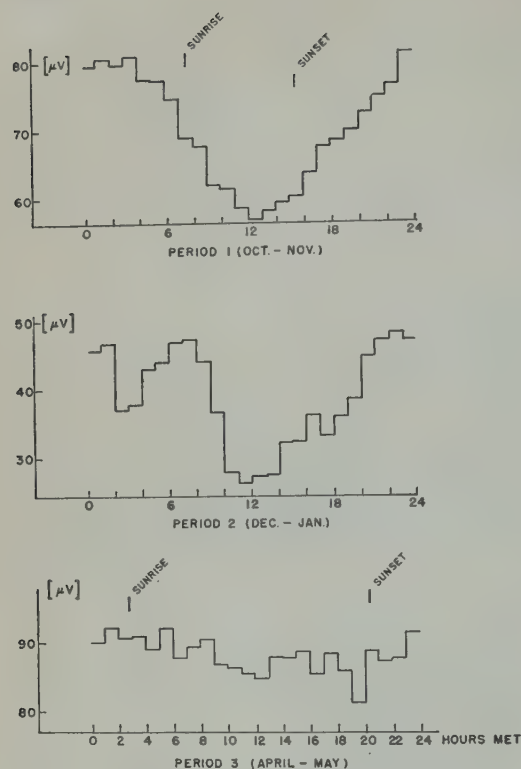


Fig. 4. Diurnal variation in the frequency range 3 to 10 kc/s.

200 to 700 c/s. For the frequency range 200 to 700 c/s the histogram corresponding to Figures 3 and 4 is shown in Figure 6. Only the December-January period has been analyzed in detail. The times of maximum and minimum intensity correspond well with those obtained for the lower- and higher-frequency range for the same period. But in this range a strong and approximately constant signal level is found between 0000 and 0600 MET, as can be seen in Figure 6.

##### 5. SEASONAL VARIATION IN INTENSITY

More data must be obtained before it is possible to draw definite conclusions of the seasonal variation of the low-frequency background noise, but some preliminary results can be found from the data analyzed.

Figure 3 indicates a relatively regular diurnal variation of the intensity in the range 20 to 200 c/s both in the winter and in the spring. The greatest mean value of the intensity in this range is found in the December-January period,

and the intensity for the April-May period is about half. As can be seen from Figure 4, a diurnal variation is observed in the range 3 to 10 kc/s in the springtime. In contrast to the lower-frequency range the maximum mean value in the 3 to 10 kc/s range is found in the December-January period, and the pattern is quite different from that of the winter results. The decrease in amplitude close to the time of sunset was not obtained in the last observed period.

From the results found in earlier investigations, it appears that the diurnal variation curve with one single maximum during the night is the most common; however, Gøssens [1956] found that the field strength (range 10 to 150 c/s) has a strong maximum between 1000 and 1700 local time and minimum during the night. The rapid recovery in amplitude a few hours before sunset for the lower-frequency range in the April-May period may be due to the sources of generating low-frequency noise and cannot be explained by the hypothesis (see section 8) that the propagation conditions determine to a large extent the low-frequency activity.

##### 6. DIURNAL AND SEASONAL VARIATION IN FREQUENCY

One parameter of the spectograms is the change in frequency of the region of maximum intensity. Thus, the peak in Figure 2 moves with frequency as a function of time and geographical conditions.

The diurnal variation in frequency of maximum intensity (frequency of maximum intensity) during the day period for the lower- and higher-frequency range is given in Figure 6. The curves show that the frequency of the highest intensity is maximum in the middle of the day with a minimum in the early morning (0600 to 0900 MET) and another late in the afternoon. No diurnal variation could be seen in the middle of the other two periods, where the curves are rather irregular.

The seasonal variation in frequency of maximum intensity is illustrated by the mean values for each period given in Table 1. As can be seen, the mean frequency in the lower-frequency range shows a maximum in the winter, in contrast to the range 3 to 10 kc/s where a spring maximum is found.



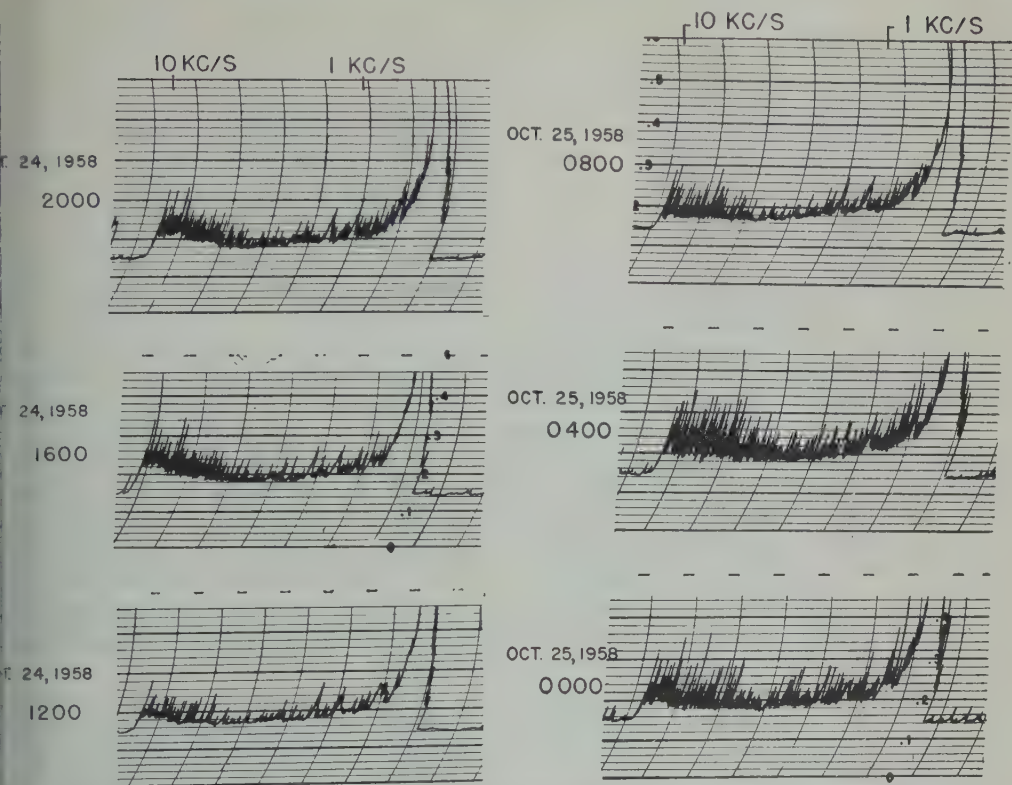


Fig. 5. Diurnal variation in the frequency range 0.5 to 10 kc/s, Oct. 24-25, 1958.

## 7. CORRELATION WITH ATMOSPHERIC DISCHARGES, MAGNETIC DISTURBANCES, AND SOLAR ACTIVITY

*Magnetic disturbances.* In an earlier paper [Gustafson, Gustafson, and Egeland, 1960] we found on eleven occasions two distinct frequency bands of several hours' duration in the low-frequency emissions. In the lower-frequency band the frequency of maximum intensity usually lies near 750 c/s, the gyro frequency for protons in the Kiruna area at an altitude of 100 km. The other band in the range 1.8 to 4.5 kc/s is most likely hiss [Watts, 1957]. Strong microbursts were shown on the Kiruna magnetogram during all the eleven observed periods of low-frequency emissions. The low-frequency emissions are quite distinct from the normal background emissions and are seen as strong enhancements in certain regions of the spectrum. The six occasions on which emission of electromagnetic energy from

protons in the low-frequency band, 600 to 950 c/s, was observed for a duration of more than two sweeps are compared in Figure 8. The length of the lines indicates the frequency widths, and the cross on each line denotes the frequency of maximum amplitude.

In order to investigate the correlation between the normal diurnal variation of the low-frequency noise and the magnetic activity, the signal level for the first period was compared with the  $Q$  indices for Kiruna (i.e., maximum deflection each quarter), which are shown in Figure 9. The mean diurnal variation of the  $Q$  indices ran relatively parallel to the first curves in Figures 3 and 4, but day by day the cross-correlation coefficient was generally found to be small (approximately 0.25 in mean value), in contrast to the findings for the 'two-noise bands' mentioned above. Strong and irregular deviations of the signal level from the normal, however, were often associated with geomagnetic disturbances.

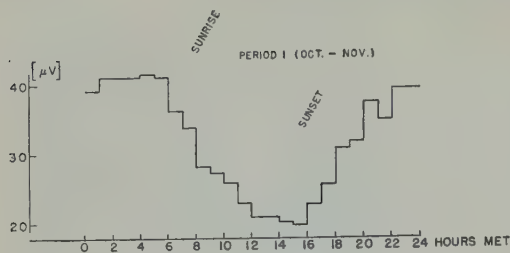


Fig. 6. Diurnal variation in the frequency range 200 to 700 c/s.

A general lack of correlation between low-frequency noise and magnetic activity has previously been noted by almost all observers; this is also indicated by the opposing intensity variations as a function of latitudes [cf. *Wilcox and Maple, 1957*]. The measurements indicated that the electromagnetic energy in the range 10 to 45 c/s is usually associated with geomagnetic disturbances. Similar strong correlation in the range 0.5 to 15 c/s was found by *Aarons and Hennisart [1953]*.

*Solar activity.* After the solar flare of importance 3+ on May 10, 1959, the frequency of maximum intensity in the range 20 to 200 c/s increased by about 30 per cent compared with the normal diurnal variation. (See Fig. 7.)

It is well known that sudden ionospheric disturbances (SID's) can affect the propagation of audio-frequency signals, but a brief investigation of SID's and low-frequency noise has yielded no direct relationship. (During an SID the electron density in the lower ionosphere is increased and its layer height is lowered considerably.)

Correlation coefficients were calculated for the 'noise' in the range 20 to 200 c/s and solar radio noise at 200 Mc/s for the April-May period. The mean value of all correlation coefficients is 0.31; the maximum value is 0.79. Therefore, there seems to be no significant correlation between solar radio noise at 200 Mc/s and the low-frequency electromagnetic radiation for this period.

*Thunderstorm activity.* It is accepted by several observers [cf. *Wilcox and Maple, 1957; Holzer and Deal, 1956*] that most, if not all, of the fluctuations in the audio-frequency range originated in thunderstorm activity. A brief examination of the comparison between low-fre-

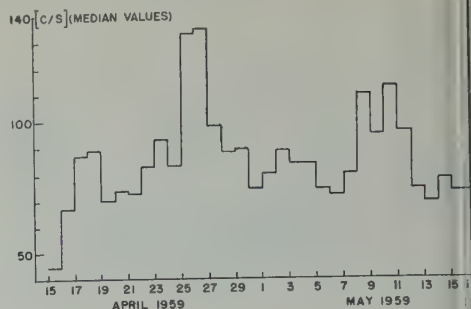


Fig. 7. Variation of frequency of maximum intensity during the period April 15, 1959, to May 15, 1959.

quency emission and the thunderstorm activity in Sweden during the April-May period yielded no correlation either in the normal variation of intensity or in the change of frequency.

## 8. PHYSICAL EXPLANATION OF THE FREQUENCY SPECTRA OBSERVED

The origin, propagation, and reflection of low-frequency electromagnetic radiation is not understood. Widely different ideas have been discussed, but little progress has been made toward explaining all types of low-frequency noise. The purpose of the following discussion is to examine briefly three possible mechanisms for low-frequency radiation, namely: (1) radiation within the earth or the ionosphere, in association with geomagnetic storms; (2) incoming electromagnetic radiation in the audio-frequency range from thunderstorm activity; (3) low-frequency radiation crossing the ionosphere from outer space.

The signal strength observed is due to the generating sources and the propagation conditions. The low-frequency electromagnetic waves may have two general modes of propagation: either they propagate along the earth's surface, via multiple reflections between

TABLE 1

| Frequency<br>Range, c/s | Mean Values, c/s, for |                 |                 |
|-------------------------|-----------------------|-----------------|-----------------|
|                         | Oct. 15–Nov. 14       | Dec. 12–Jan. 11 | April 15–May 14 |
| 20–200                  | 71                    | 123             | 86              |
| 200–700                 | 281                   |                 |                 |
| 3000–10,000             | 7630                  | 8420            | 8450            |

and the ionosphere, and directly (for distances longer than 1000 km, the ground wave negligible); or else they propagate essentially along the lines of force of the geomagnetic field, analogously with whistlers.

None of the sources mentioned can by itself explain the low-frequency noise observed. Perhaps the results could be explained by contributions from the three different hypotheses together with the propagation conditions. The hypotheses and the propagation conditions are described in more detail in the following paragraphs.

*Terrestrial origin.* Ionospheric currents in the earth currents, and the magnetic effects of these two current systems accumulate and may produce a great part of the noise observed, especially in the lowest frequency range. The association between low-frequency radiation in the range 10 to 45 c/s may thus be of ionospheric origin.

*Thunderstorm activity.* It is assumed by several observers that the fluctuations in the audio-frequency range originate in thunderstorm activity and reach the receivers by traveling between the ground and the ionosphere. This assumption is based on the following reasons: (1) The signal strength in the range of about 50 to 100 c/s decreases with increasing latitudes in the same way as the thunderstorm activity. (2) Far-by lightning discharges cause a marked increase in the signal level, in contrast to our findings. (3) Little correlation with magnetic disturbances is found.

In the absence of local discharges, the observed frequency spectra are explained in terms of electromagnetic signals generated in regions of thunderstorm activity and propagated poleward. The signal intensity, with a strong nighttime maximum and a daytime minimum, is explained in part by the different propagation

conditions.<sup>1</sup> (The propagation factor for the waves in this range is about twice as great in the winter as in the summer [Bowe, 1953], which corresponds well with the results obtained in this investigation in the range 20 to 200 c/s.)

Departures from the diurnal variation described were found by Goldberg [1956], but he also described the 'noise' as largely due to lightning activity.

These hypotheses cannot explain all the important features of the low-frequency activity observed, for example, frequency lines, bands, and trains in the observed data.

*Extraterrestrial origin.* One possible source of low-frequency radiation is the emission due to protons revolving around the geomagnetic field lines at the gyro frequency [Aarons, Gustafsson, and Egeland, 1960]. Another possibility is the plasma oscillations of free protons in the ionosphere [Allcock, 1957].

In one theoretical study, Gallet and Helliwell [1959] assume that a part of the low-frequency radiation is generated in the earth's exosphere. Selective traveling-wave amplification in the outer ionosphere is postulated. The input signal is assumed to be provided by whistler energy, Cerenkov radiation, and thermal radiation. The low-frequency noise is excited by streams and bunches of high-speed particles precipitating into the ionized atmosphere in the presence of the earth's magnetic field and propagated in the same manner as whistlers [Gallet, 1959].

This theory can explain discrete tones, bands, and trains of low-frequency activity. Such low-

<sup>1</sup> In spite of this simple geometrical picture, the theoretical treatment of the problem of low-frequency electromagnetic waves is especially difficult, and a detailed solution of this problem has been possible only in some special, idealized cases, such as mode theory, approximated geometrical-optics theory, and Pöeverlein's theory.





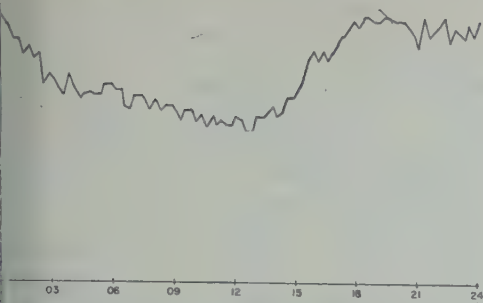


Fig. 9. The mean  $Q$  indices curve for Oct. 12 to Nov. 11, 1958.

frequency noise is usually very well correlated with solar and magnetic activity, and not associated with lightning discharges.

Willis [1957] states that the Cerenkov radiation contributes to the signal intensity and is the cause of 4 to 5 kc/s auroral signals. Aarons [1954] compares the gas discharge mechanism with the aurora; he first discusses the possibility of detecting the protons and other ions by emission.

## 9. DISCUSSION

The findings discussed here do not solve the problem whether the electromagnetic radiation in the audio-frequency range is of terrestrial origin (i.e., thunderstorm activity or ionospheric) or is due to extraterrestrial sources. The Cerenkov radio emission, an analogy to traveling-wave tube amplifiers, and gas discharge mechanism) and strongly correlated with auroral activity.

As far as we know, almost all investigations of 'noise' in the range 5 to 10,000 cps [cf. Willis and Maple, 1957, and Holzer and Deal, 1956] have resulted in the acceptance of the explanation that the fluctuations in the audio-frequency range originated in thunderstorm activity. An exception is Gallet and Helliwell [1959], who postulated that at least a large portion of this noise must arise in the earth's outer ionosphere. As observed in the April-May 1956, lightning discharges from 1000-km distance increase the signal levels, and hiss and noise during aurora were strongly associated with magnetic activity. Therefore, it seems reasonable to conclude that none of the sources mentioned can alone explain the low-frequency

activity observed. The most probable explanation would be a contribution from the different sources.

It would be expected that, if the origin of the background noise were geomagnetic, an increase in level would be observed at Kiruna in comparison with lower latitudes. Just such a tendency is observed when Kiruna data in the range 20 to 200 c/s are compared with data taken at 32°N. Since the equipment used was somewhat different, however, absolute calibrations are difficult.

In the absence of local thunderstorm activity (within 1000 km) at high latitudes, especially in the winter, the signal levels are explained in terms of radiation from thunderstorm activity in the region of the equator, and a close correlation is very difficult to obtain on account of the sources and the propagation conditions.

The general spectrum observed is not dissimilar to that which would be expected from the observations of individual atmospheric layers. The slow tail maximum at 100 c/s is seen in these observations, as is the absorption band centered at 2.5 kc/s and a general increase in signal level to 10 kc/s. The general appearance of the 3 to 10 kc/s band observed in these experiments indicates that the maximum frequency is somewhat above 10 kc/s.

Since the diurnal variations of the low-frequency signals are quite similar in experiments performed at different longitudes [Aarons, 1956, and this study], if the origin is world-wide thunderstorms, the propagation mechanism is certainly effective in changing the amplitudes of the signals received.

## 10. SUMMARY

The background signals recorded show a similar variation throughout the entire spectrum studied. A minimum signal is received during the day and a maximum at night. It seems reasonable to conclude that this is an absorption effect of the daytime ionosphere, since the diurnal pattern is somewhat similar over the entire frequency band. Layer heights also affect the signals, since the ratio of night to day signals varies as a function of frequency in a rather complex manner.

The frequency of the maximum signal also has a diurnal variation. In the range from 30 to 600 c/s, nighttime frequencies are higher than

daytime in the December-January period, but the reverse is true in the October-November period.

The December-January period has a higher frequency of maximum intensity in this part of the spectrum than the other periods, indicating perhaps the influence of the large zenith distance of the sun during this period (Kiruna is at high latitude).

The origin of the signals is not clearcut; atmospheric from great distances probably are the most important component, with background magnetic fluctuations also contributing to the lower band.

*Acknowledgments.* We should like to acknowledge the assistance and encouragement of Dr. B. Hultqvist, Director of the Kiruna Geophysical Observatory. Mr. William Barron of the Air Force Cambridge Research Center assisted in the initial operation and organization of the program. The work was partially supported by the EO-ARDC under contract AF61(514)-1314.

#### REFERENCES

- Aarons, J., Doctoral thesis, University of Paris, June 1954.  
 Aarons, J., *J. Geophys. Research*, **61**, 647, 1956.

- Aarons, J., G. Gustafsson, and A. Egeland, *Nature*, **185**, 148, 1960.  
 Aarons, J., and M. Henissart, *Nature*, **172**, 1953.  
 Allcock, G. McK., *Australian J. Phys.*, **10**, 1957.  
 Bowe, P. W. A., *Phil. Mag.* **44**, 833, 1953.  
 Egeland, A., *Inst. Theoret. Astrophys. Rep.*, Blindern, Oslo, 1959.  
 Ellis, G. R. A., *J. Atmospheric and Terrest. Phys.*, **10**, 302, 1957.  
 Ellis, G. R. A., *J. Planetary and Space Sci.*, **1**, 1959.  
 Gallet, R. M., *Proc. IRE*, **47**, 211, 1959.  
 Gallet, R. M., and R. A. Helliwell, *J. Res. NBS*, **63D**, 21, 1959.  
 Goldberg, P. A., *Nature*, **177**, 1219, 1956.  
 Holzer, R. E., and O. E. Deal, *Nature*, **177**, 1956.  
 Maple, E., *Trans. Am. Geophys. Union*, **38**, 1957.  
 Menzel, D. H., and W. W. Salisbury, *Nature*, **91**, 1948.  
 Watts, J. M., *J. Geophys. Research*, **62**, 199, 1957.  
 Wilcox, J. B., and E. Maple, *Navord Rept.*, U. S. Naval Ordnance Laboratory, White Sands, Md., July 9, 1957.  
 Willis, H. F., *Nature*, **161**, 887, 1948.

(Manuscript received June 9, 1960.)



## Correlation of an Auroral Arc and a Subvisible Monochromatic 6300 Å Arc with Outer-Zone Radiation on November 28, 1959

B. J. O'BRIEN AND J. A. VAN ALLEN

*Department of Physics and Astronomy, State University of Iowa  
Iowa City, Iowa*

F. E. ROACH

*Boulder Laboratories, National Bureau of Standards  
Boulder, Colorado*

AND

C. W. GARTLEIN

*IGY World Data Center A, Cornell University  
Ithaca, New York*

**Abstract.** During a severe geomagnetic storm on November 28, 1959, two Geiger tubes on satellite Explorer VII (1959c) found anomalies in the outer radiation zone at an altitude of about 1000 km which appear to be correlated in space and time with optical emissions from the atmosphere beneath. Very intense narrow zones of radiation were detected over a visible aurora during one pass. The radiation in three such zones was harder toward low latitudes. On three subsequent passes the radiation zone was deduced to be over a subvisible 6300 Å arc, whose brightness diminished as the radiation zones became less intense. The correlation is discussed.

**Observations and discussion.** Following a sudden commencement at 2350 Z on November 27, 1959, active auroras were visible at low latitudes in North America, and a subvisible monochromatic 6300 Å arc [Barbier, 1958; Roach and Provich, 1959a, b] was detected by photometers at Cactus Peak, Fritz Peak, and Sacramento Peak throughout the local night of November 27-28. Two Geiger tubes on the satellite Explorer VII detected anomalies in the pattern of particle radiation in the outer zone [Van Allen, 1959] over regions of these optical emissions. The relationship of these phenomena is discussed in this note.

Visual observations and all-sky-camera photographs of the auroras were made under the general supervision of the auroral group at Cornell University.

Photometric observations of the subvisible red arc were made by three groups at Cactus Peak, Fritz Peak, and Sacramento Peak. By triangulation

methods, the height of the red arc throughout the night was found to be 400 km (Roach, Moore, Cronin, and Silverman, 'The height of a monochromatic auroral arc,' article to appear in this journal).

Outer-zone observations were made with two Geiger tubes on the satellite Explorer VII, which has an apogee altitude around 1085 km and perigee altitude around 555 km. The apparatus has been described in detail by Ludwig and Whelpley [1960]. One Geiger tube, the Anton 302, responds with low efficiency to bremsstrahlung from electrons above about 30 keV; the other, a lead-shielded Anton 112, responds to those above about 70 keV. In an omnidirectional intensity of penetrating charged particles of  $1/\text{cm}^2 \text{ sec}$ , the counting rates are respectively 0.54/sec with a scaling factor of 2048 and 7.1/sec with a scaling factor of 128.

In Figure 1 is shown the counting rate versus time of each tube as Explorer VII traversed

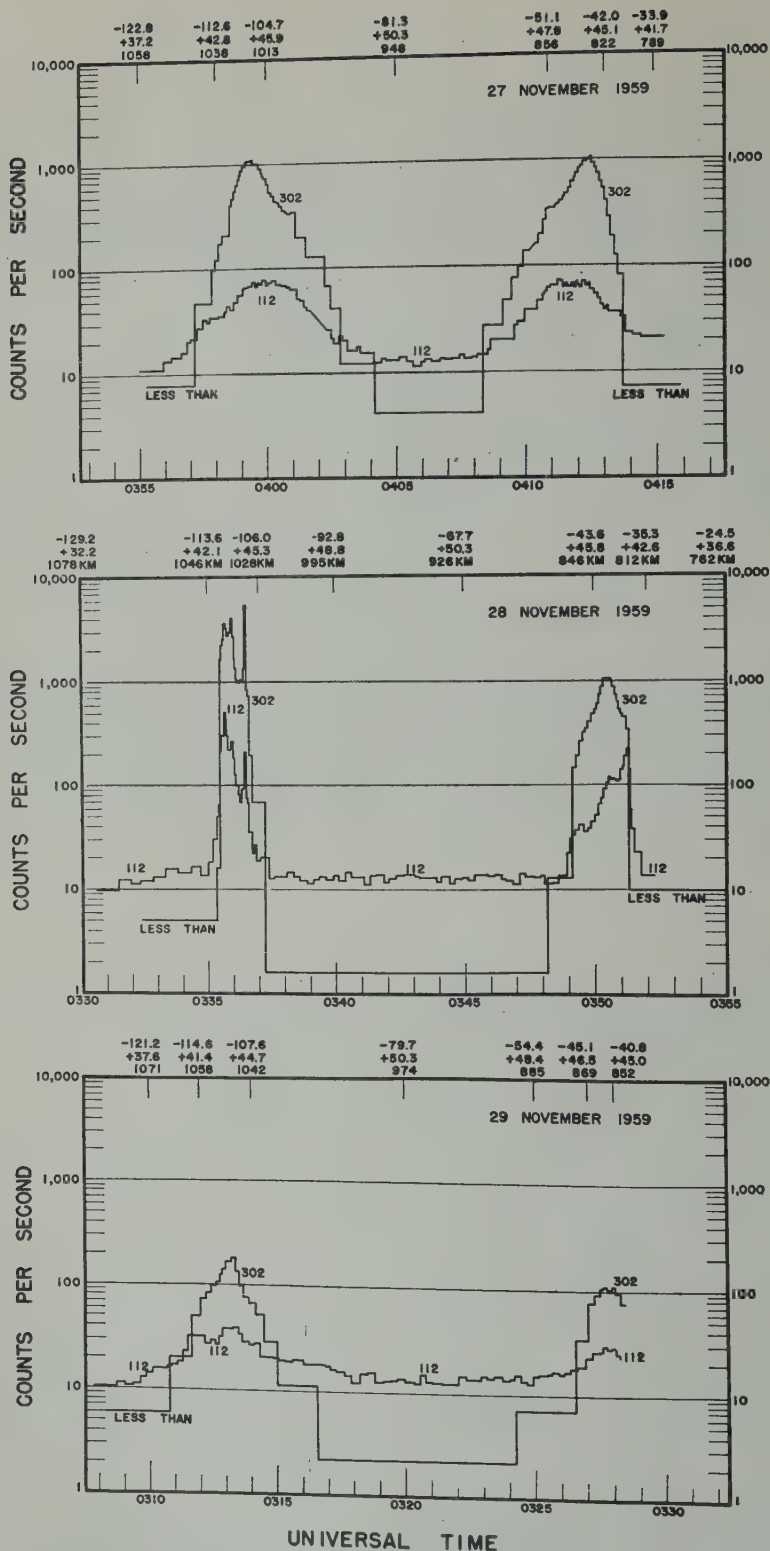


Fig. 1. True counting-rates of 302 and 112 Geiger tubes on Explorer VII on passes in roughly the same positions in space on November 27-29, 1959. The graphs are aligned by the time at which the satellite crossed geographical latitude  $40^{\circ}$  N. The abscissae of all three graphs indicate universal time.

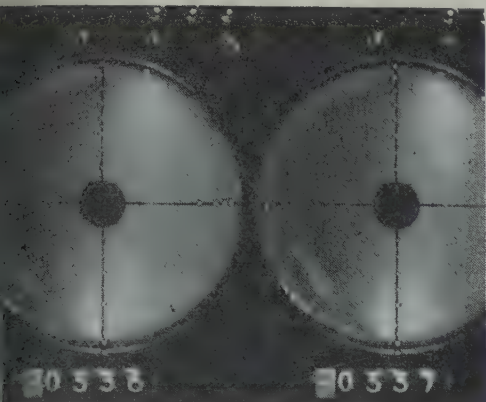


Fig. 2. All-sky-camera photograph at 0336 Z on November 28, 1959, at Choteau, Montana. Geographic north is toward the right, and east is at the bottom of the photograph. The lights on the north-south axis represent the respective positions of the auroras at 100-km altitude and at progressively  $0.5^\circ$  through  $4.5^\circ$  of latitude from the camera.

outer zone on its northbound pass and then passed through it again on its southbound pass from 0355 to 0415 Z on November 27, 1959, when the radiation zone was relatively stable. The pattern of the corresponding pass on November 28 in which nearly the same position in space is also shown, and it is apparent that marked changes have occurred. Finally, to complete the sequence, the corresponding pass on November 29 is shown. A notable feature of this figure is the very intense peak of radiation shown at 0336:30 Z on November 29, 1959. Among many hundreds of examples of the outer zone seen by Explorer VII, this peak is unique in two respects: it is the narrowest peak seen to date (traversed by the satellite in 3 seconds; it contains the softest radiation seen to date, with a ratio of 302 to 112 of 27 to 1, compared with the usual outer-zone value of around 14 to 1.

In Figure 2 are reproduced two all-sky-camera photographs taken at 0336 and 0337 Z on November 28 at Choteau, Montana, by John G. Gibson, showing a broad auroral arc in the north. Fainter aurora extends at least  $0.5^\circ$  to the south, and a diffuse medium-intensity arc is seen on the northern side. A visual observer, G. Schmidt of Browning, Montana, also observed this feature at 0345 Z. This auroral event began at about 0300 Z and by 0335 Z had

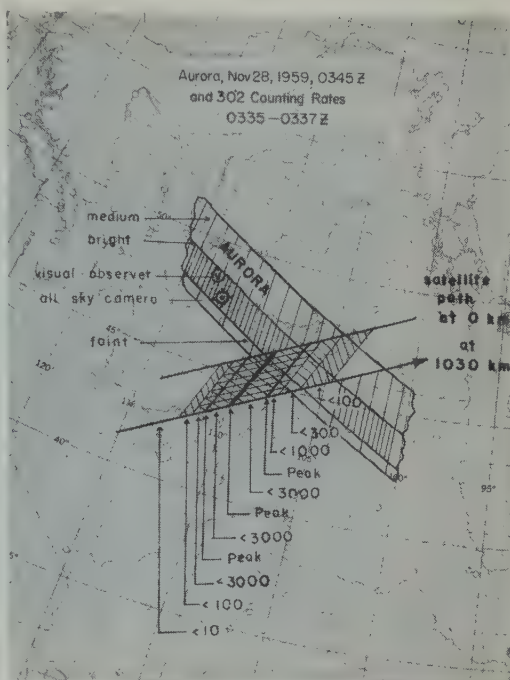


Fig. 3. Positional comparison of the auroral arc at an unknown height and the satellite radiation observations around 0336 Z on November 28, 1959. Three heavy lines show projection of counting-rate peaks from the satellite orbit downward along field lines to the ground. The symbol  $<3000$  indicates a counting rate between 1000 and 3000/sec,  $<1000$  being between 300 and 1000/sec,  $<300$  being between 100 and 300/sec,  $<100$  being between 100 and 10/sec.

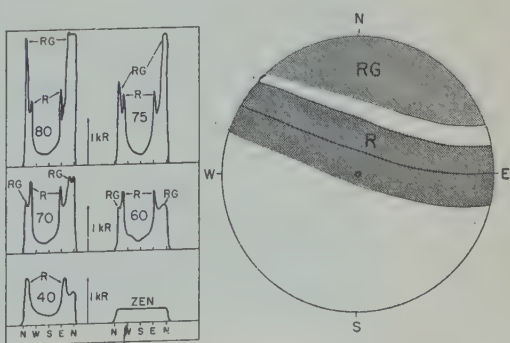


Fig. 4. Photometer data at 0855 Z from Fritz Peak, and the deduced geographic location of the 6300 Å arc (R) and the auroral activity (RG) which consists of 6300 and 5577 Å.



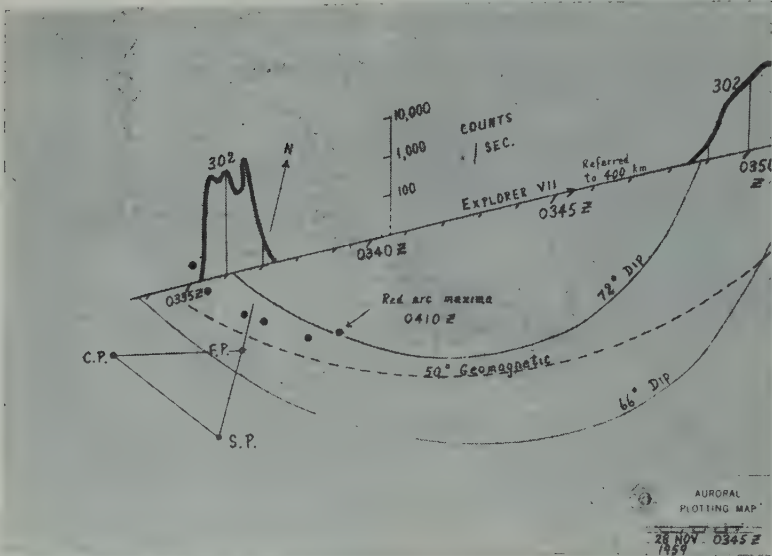


Fig. 5a

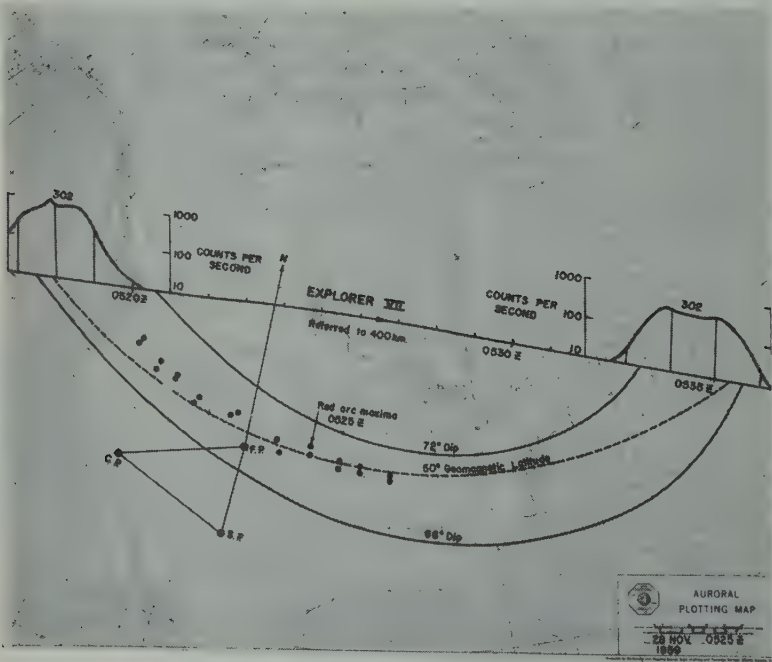


Fig. 5b

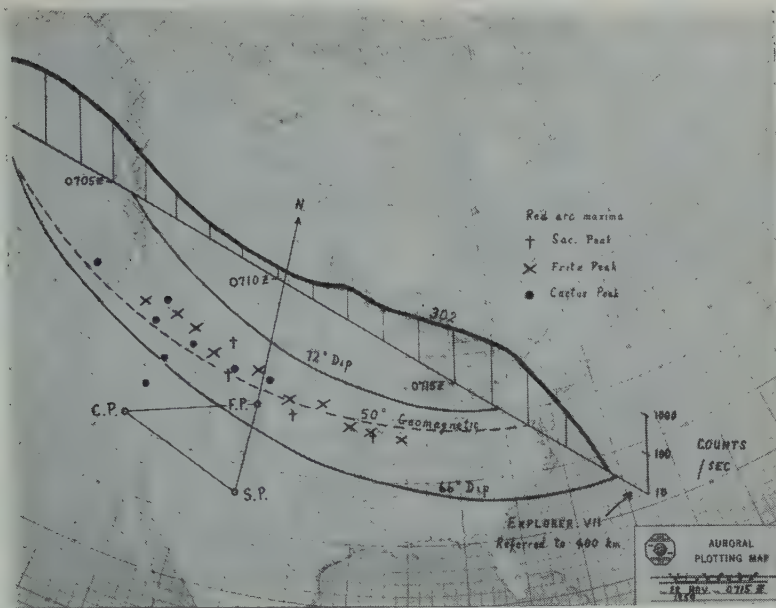


Fig. 5c

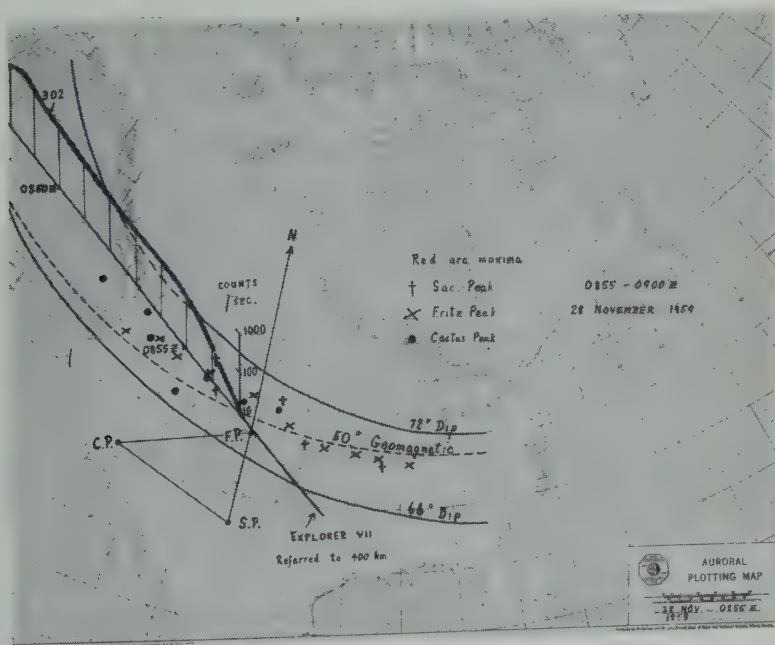


Fig. 5d

Fig. 5. Positions of red-arc maxima and satellite data (the latter projected downward to 400-km altitude) at several times on November 28, 1959. The 302 counting rate is sketched on a logarithmic scale on the locus. The centered-dipole geomagnetic latitude of 50° is marked, as are two dip lines roughly defining the width of the red arc.

moved southward about  $3.5^\circ$  of latitude. It continued to move southward another  $1.5^\circ$  until it changed into ray structure at about 0455 Z.

In Figure 3 these broad auroral forms are plotted from the all-sky photograph to show ground positions of overhead auroras. The visual observer delineated the lower edge of the bright arc. His data correspond to the junction of bright and medium aurora. The subsatellite locus is plotted, and this path is projected at right angles to the isocline down the magnetic field lines to give the path projected on the ground. The hatching between the projection lines shows the radiation intensity measured by the 302 tube. The three peaks exceed 3000 counts/sec (Fig. 1). The downward projections of the anomalous structure lie over the auroral forms. Hence, if charged particles detected by the satellite in the structure had continued spiraling down the magnetic field lines they would have coincided in space and time with the aurora.

More detailed discussion of the relationship between the satellite-observed particles and the visible aurora is necessarily tentative, owing to inadequate knowledge of the spectrum and altitude dependence of intensity of the former. Simple energetic considerations show that the qualitative correlation established above also has quantitative plausibility. This may be seen as follows. On the basis of present knowledge it is reasonable to suppose that the particles that are being detected by the satellite equipment are electrons whose number-energy spectrum rises steeply toward low energies. For a typical particle energy of 50 keV, it is found by laboratory measurement that the counting rate of the 302 must be multiplied by a factor of the order of  $10^7$  to  $10^8$  to obtain the omnidirectional intensity of electrons. Hence, the energy flux at the position of the satellite is likely to be of the order of  $10^4$  ergs/cm<sup>2</sup> sec (or greater), or some  $5 \times 10^5$  times the optical energy emitted in the 5577 Å line in an easily visible aurora of 10 kilorayleigh brightness (International brightness coefficient II) [Hunten, Roach, and Chamberlain, 1956]. McIlwain [1960] has found from rocket flights into auroras over Fort Churchill that the gross efficiency of excitation of atmospheric emissions in the visible region by electrons of similar energy is about 0.2 per cent. Hence the energy flux in the peak is several hundred times

that required to excite the observed aurora. Indeed, this factor may be considerably greater on account of the neglect of electrons below the detection threshold of the satellite-borne counter. From these considerations it is not to be plausible that the particles in question are in quasi-trapped orbits in the geomagnetic field and are absorbed in the atmosphere after hundreds of latitudinal oscillations (i.e., in a period of time of the order of tens of minutes). (The replenishment of low-mirror-altitude orbits is neglected.) Furthermore, the temporal stability of the arc is another indication that the aurora was not excited by an instantaneous dumping of intense particle flux.

One further point to be noted in Figure 1 is the marked increase in hardness toward low latitude of the radiation in the three peaks around 0400 Z on November 28. The significance of this feature is not immediately apparent.

The other optical emission of special relevance to this study was the wide (about  $6^\circ$  to  $8^\circ$  latitude) 6300 Å arc which was detected from 0410 through 1000 Z. The arc moved slowly southward, showed gradual changes in peak intensity, and was at a height of 400 km (Roach, Moore, Cronin, and Silverman, article to appear in this journal). It was quite extensive in longitude, following approximately along line of constant dip angle, as has been noted with similar arcs previously [Roach and Marovich, 1959a, b]. The photometer data showing the red arc from 0855 Z from Fritz Peak are depicted in Figure 5.

In Figure 5a, b, c, and d the positions of peak intensity of the red arc are shown for various times when Explorer VII was passing over North America. The latitude thickness of the red arc is indicated roughly by the two lines of constant dip for the 0525, 0715, and 0855 Z maps, and the same lines are drawn for the 0345 Z map although the red arc was farther north at that time.

The pertinent satellite path on each map is projected along magnetic lines to the red-arc altitude of 400 km is shown on these maps, together with a logarithmic sketch of the counting rate during each pass.

The first photometric observation at 0410 Z was made after the satellite had passed at 0345 Z, at a period when the arc was moving steadily south. Although some spatial correlation of



is indicated, the time difference prevents conclusions, and the data are included here for completeness only.

In the other three maps the photometric and satellite observations were made nearly simultaneously. Plausible extrapolation in longitude of the red arcs on the 0525 and 0715 *Z* maps shows striking positional correlation with the satellite observations. The two classes of observations shown in the 0855 *Z* map were coincident in both space and time and provide even more convincing evidence of their relationship.

It should also be noted that both the peak ionizing rates and the red-arc intensity decreased from pass to pass.

In view of the similar locations and width of the red arc and the outer zone defined by the 302 center, and in view of the gradual decrease in intensity of each, we believe that the two are associated. In particular, it would appear that the gradual depletion of the outer zone is due to progressive atmospheric scattering-down and absorption of trapped particles, and that the corresponding energy influx is the source of energy for the excitation of the 6300 Å radiation.

By an argument similar to that in an earlier section it is estimated that the presumed influx of energy from the detected particles could have maintained a 1-kR intensity 6300 Å arc for a period of the order of 10 hours only if the efficiency of excitation is several per cent.

Observations by Arnoldy, Hoffman, and Finkler [1960] with Explorer VI show that there was a very great loss of trapped radiation near the equatorial plane after a sudden commencement in August 1959. As the August event seemed similar in nature to the November event which is the subject of the present paper, this evidence may be cited to support our belief that the decreasing intensity at altitudes of about 1000 km is due to the absorption of these particles in the atmosphere and is not due to the more conceivable alternative—namely, gradual outward drift of their mirror points to an altitude above that of the satellite observations.

Many problems remain unanswered in this study. For example, it must be borne in mind that on this occasion we have no knowledge of the fluxes of electrons below around 30 kev. Furthermore, we have insufficient knowledge of the efficiency of energy conversion from electron

kinetic energy into optical emissions.<sup>1</sup> One further considerable problem (of long standing) lies in the fact that in the red arc little or no 5577 Å radiation was emitted, even though the energy required to raise atomic oxygen to the parent level for this line is only 2 ev more than to lift it to the parent level for 6300 Å.

The correlation of the red arc with the aurora under the radiation peak at 0336 *Z* is also uncertain. The intensity of a few kilorayleighs of 6300 Å would be insufficient to register either on the human eye or the all-sky-camera photographs. Previous observations [Roach and Marovich, 1959b] showed conclusively that a 6300 Å arc was not merely above a green auroral feature, but indeed on the low-latitude fringe of an aurora containing 5577 Å (i.e., geographically distinct from it). If one assumes that the 6300 Å arc on November 28, 1959, was of a similar type, one would suppose that the *RG* of Figure 4 was seen by the visual observer and the all-sky camera whereas the red arc *R* was not.

*Concluding remarks.* Since the early work on the trapped radiation it has been clear that the particles present in the outer zone have the potential capability of producing bright atmospheric emissions. Moreover, the lower horn of the intense part of the outer zone 'hovers' just above the appreciable atmosphere. The evidence presented in this paper constitutes the direct observation of an actual instance of such a 'dumping' process, under magnetically disturbed conditions. It is hoped to extend and refine knowledge of the details of the dumping mechanism in subsequent work and to study the atmospheric emissions excited by the usual, quiescent influx of trapped particles from the outer zone.

*Acknowledgment.* The satellite part of the observational work was supported by the National Aeronautics and Space Administration.

<sup>1</sup> We do not mean to imply that the 6300 Å emission was directly excited on bombardment of atmospheric oxygen by energetic electrons. King and Roach (in an article to be submitted to *J. Research NBS*) have advanced a hypothesis that the role of the dumped electrons may have been to raise the local temperature and mix the atmospheric constituents, thereby starting a suggested chemical reaction that yielded 6300 Å.

## REFERENCES

- Arnoldy, R. L., R. A. Hoffman, and J. R. Winckler, Observations of the Van Allen radiation regions during August and September 1959, *J. Geophys. Research*, *65*, 1361, 1960.
- Barbier, D., l'Activité aurorale aux basses latitudes, *Ann. géophys.*, *14*, 334, 1958.
- Hunten, D. M., F. E. Roach, and J. W. Chamberlain, A photometric unit for the airglow and aurora (and references therein to work of Rayleigh and Seaton), *J. Atmospheric and Terrest. Phys.*, *8*, 345, 1956.
- Ludwig, G., and W. A. Whelpley, Corpuscular radiation experiment of satellite 1959 (Explorer VII), *J. Geophys. Research*, *65*, 1119, 1960.
- McIlwain, C. E., Direct measurement of particles producing visible auroras, *State Univ. Iowa research Rept. 60-7*, 59 pp., June 1960; and *Geophys. Research*, *65*, 2727, 1960.
- Roach, F. E., and E. Marovich, A monochromatic low latitude aurora, *J. Research NBS*, *63D*, 1959a.
- Roach, F. E., and E. Marovich, Aurora of October 22/23, 1958, at Rapid City, South Dakota, *Research NBS*, *64D*, 205, 1959b.
- Van Allen, J. A., The geomagnetically trapped corpuscular radiation (and references therein to primary sources), *J. Geophys. Research*, *64*, 1683, 1959.

(Manuscript received July 1, 1960.)

## *F*-Region Traveling Disturbances and Sporadic-*E* Ionization

L. H. HEISLER AND J. D. WHITEHEAD

*Radio Research Board Laboratory, Electrical Engineering Department  
University of Sydney  
Sydney, Australia*

**Abstract.** Ionosonde records taken at Sydney (33°52'S, 151°11'E) during July 1955 have been analyzed statistically for the occurrence of *F*-region traveling ionospheric disturbances and *E<sub>s</sub>*. It is shown beyond reasonable doubt that the phenomena are associated, and that, as previously suggested by one of the authors, the advent of a traveling ionospheric disturbance is sometimes accompanied by an increase in *E<sub>s</sub>* critical frequency. An apparent small secondary effect in *E<sub>s</sub>* occurring some minutes after the *F* anomaly is also discussed.

**Introduction.** For some time it has been suggested that a correlation may exist between region traveling ionospheric disturbances (TID's) and sporadic-*E* ionization. Thomas [1958] suggests that the passage of a sporadic cloud should be associated with some disturbance in the *F* region, and Bibl and Rawer [1959] discuss the generation of a sporadic-*E* layer of the sequential type following *F*-region disturbances. Previously Munro [1950] had tried to find any association between *E<sub>s</sub>* and TID's observed on a fixed frequency of 5.8 Mc/s. This type of disturbance is small in magnitude, however, and recently Heisler [1958] has described large TID's which produce gross distortions on ionosonde records, suggesting furthermore [Heisler 1959] that these are related to sporadic-*E* ionization. Specific examples are quoted to show that the critical frequency *E<sub>s</sub>* increased at a series of stations during the passage of a traveling disturbance in the *F* region. Since specific examples of events in the ionosphere were considered an insufficient basis on which to establish a theory correlating the phenomena, statistical study of the occurrence of *E<sub>s</sub>* and TID's was undertaken, to test the hypothesis. The results presented here not only show that *F*-region disturbances and sporadic *E* are associated beyond reasonable doubt but also reveal several other interesting features.

**Analysis.** The records analyzed consist of ionosonde recordings taken every minute during daylight hours in July 1955, July being a month when *E<sub>s</sub>* occurs infrequently and only rarely penetrates the *F* region. Furthermore, TID's dur-

ing July are mainly the cusp-type anomalies, which are easily recognized and whose time of occurrence can be placed within 5 minutes. The main features of this anomaly, described by Munro and Heisler [1956], consist of a reduction in critical frequency of the *F<sub>2</sub>* region followed by the appearance of a cusp at frequencies near *f<sub>o</sub>F<sub>2</sub>*, and at a high virtual height. This cusp progressively falls in height and shows a resultant complexity in the main *F<sub>2</sub>* and *F<sub>1</sub>* traces. Heisler [1958] has already pointed out that because of the transient nature of TID's the usual ionosonde record taken every 10 minutes is too infrequent to allow satisfactory analysis of the phenomena. Since the appearance of the cusp is the most prominent feature of a traveling disturbance, we have defined the time of occurrence of the TID's in the analysis as the time at which the cusp first appears, which is about 3 minutes after *f<sub>o</sub>F<sub>2</sub>* starts to decrease. A total of 330 TID's was found from the month's records.

As the analysis of just 1 month's records involved inspection of 20,000 individual ionosonde records, it was decided to allot an *E<sub>s</sub>* index to each record on the basis given in Table 1, rather than noting *f<sub>o</sub>E<sub>s</sub>*.

The superimposed epoch method was used to determine the variation in *E<sub>s</sub>* index during a traveling disturbance, account being taken both of the diurnal variations of frequency of disturbances and *E<sub>s</sub>* index and of the fact that TID's do not occur at random but are less likely to occur immediately before and after another one. The remaining variation *f*(*t*) is shown in



TABLE 1

| $E_s$ Index | Record                            |
|-------------|-----------------------------------|
| 0           | No $E_s$                          |
| 1           | Sign of $E_s$                     |
| 2           | Definite $E_s$                    |
| 3           | $E_s$ trace (o or x) above 4 Mc/s |
| 4           | Blanketing $E_s$                  |

Figure 1. The points lie on a deceptively smooth curve, owing to the autocorrelation of  $E_s$  index over the period. The variations in  $f(t)$  are of the order of 0.05, whereas the standard deviation of the mean values (calculated to take account of this correlation) is 0.027; thus  $f(t)$  is not inconsistent with a random variation of  $E_s$ .

However,  $f(t)$  does show quite a rapid change during the TID's, and it is interesting to see whether  $df/dt$  has values that could not be accounted for by a purely random variation. The mean values,  $df/dt$ , are plotted in Figure 2, together with its standard deviation. According to the  $\chi^2$  test of significance, the probability that these values of  $df/dt$  are due to a random distribution with the standard deviation is less than 1 per cent. The largest peaks and dips of Figure 2, therefore, represent actual variations, and the diagram shows that 13 minutes after a traveling ionospheric disturbance the  $E_s$  index reaches a maximum value.

A definite association between the two phenomena having been demonstrated, it is interesting to see whether the 'damped oscillation' of Figures 1 and 2 is real or due to random variations. During the analysis it became obvious that many TID's showed no  $E_s$  at all. The first step, therefore, is to select disturbances that have a significant  $E_s$  relationship. For each disturbance the 'significance'

$$S = \int_0^{+20 \text{ min}} (df/dt)(\overline{df/dt}) dt$$

was found, the integral being taken over the first cycle of oscillation of  $df/dt$  only. If disturbances having a large positive value of  $S$  also show the remaining cycles of oscillation, these may represent a genuine effect. Figure 3 gives the histogram of values of  $S$  and also a Gaussian error curve fitted to values of  $S$  greater than zero.

It is seen that several of the disturbances show much larger positive values of  $S$  than would be expected; out of the 330 TID's, about 200 appear to be followed by an increase in sporadic  $E$  ionization. We now consider these particular disturbances. In Figure 4 is plotted the mean value of  $df/dt$  for those disturbances for which  $S > 1.42 \times 10^{-8}$  (indices)<sup>2</sup> (min)<sup>-1</sup>. It is seen that the 'damped oscillation' appearance of Figures 1 and 2 has almost completely disappeared. A small negative (but not significant) correlation was found between disturbances having

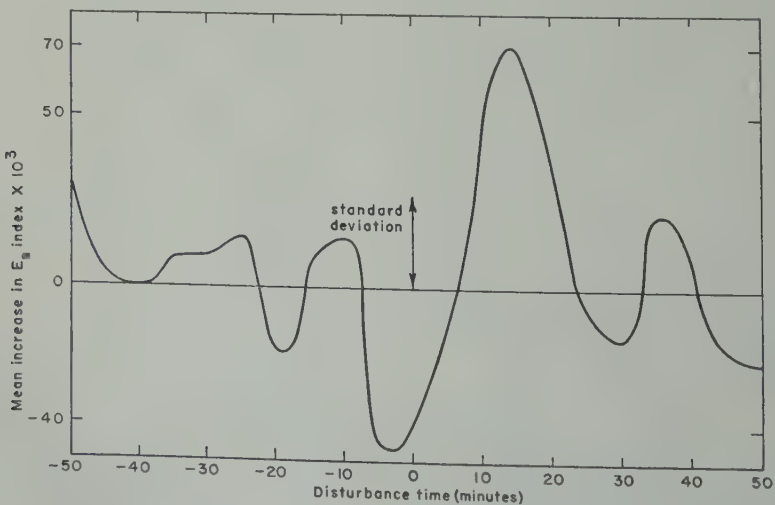


Fig. 1. Function  $f(t)$ , the remaining variation in the  $E_s$  index with disturbance time.

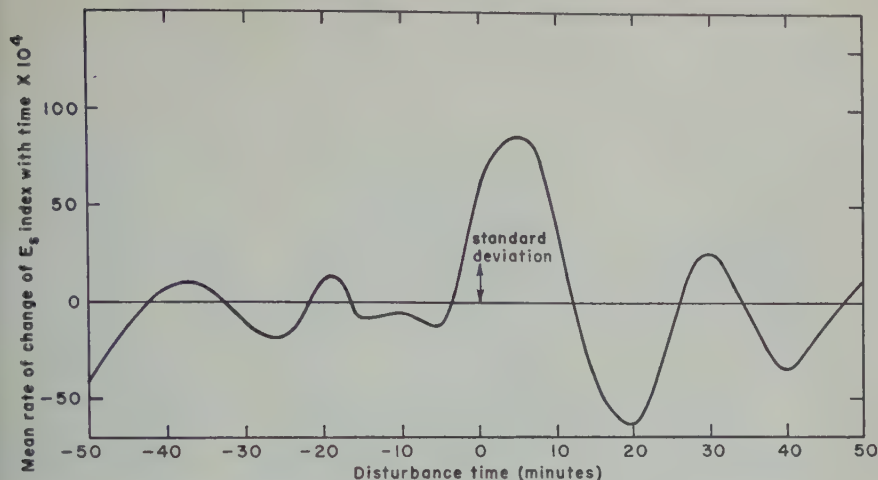


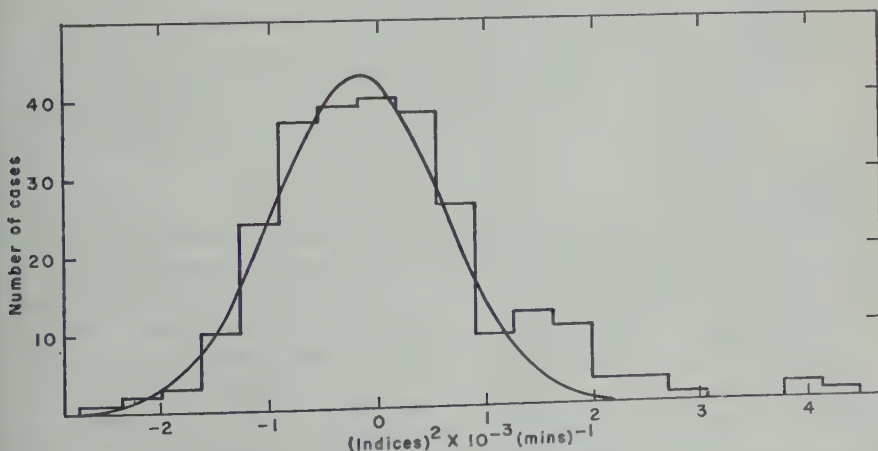
Fig. 2. Function  $df/dt$ , the rate of change of  $E_s$  index with disturbance time.

ative value of  $S$  and those showing the sec-  
cycle of the oscillation (30 to 40 minutes  
r the disturbance time).

It was concluded from this analysis that ap-  
proximately 1 TID in 15 showed an  $E_s$  effect.  
For the disturbance time, the  $E_s$  index starts  
increase, reaches a maximum of about unity  
representing an approximate increase in  $f_oE_s$  of  
10 Mc/s some 13 minutes later, and returns to  
normal 25 minutes after the disturbance time.  
A search was made for a diurnal variation in  
mean value of  $S$  to see whether a TID was  
more likely to be correlated with an  $E_s$  effect  
particular times of day. The variation, shown

in Figure 5, suggests peaks at 0900 and 1300  
hours local time, though the errors are too large  
to make this result significant. A comparison  
with the diurnal variations of  $E_s$  index and  
traveling disturbance frequency (also illustrated  
in Fig. 5) shows that all three have peaks  
around these times.

*Examination of individual disturbances.* It  
has been shown above that certain individual  
disturbances have a marked  $E_s$  correlation. A  
typical example of such a disturbance, which  
has a significance figure of  $4.18 \times 10^{-8}$  (in-  
dices)<sup>2</sup> (min)<sup>-1</sup>, is given in Figure 6: tracings of  
an ionosonde record at Camden (150°40'E,



3. Histogram of values of the significance  $S$ , and the Gaussian error curve fitted to values of  $S$  less than zero.

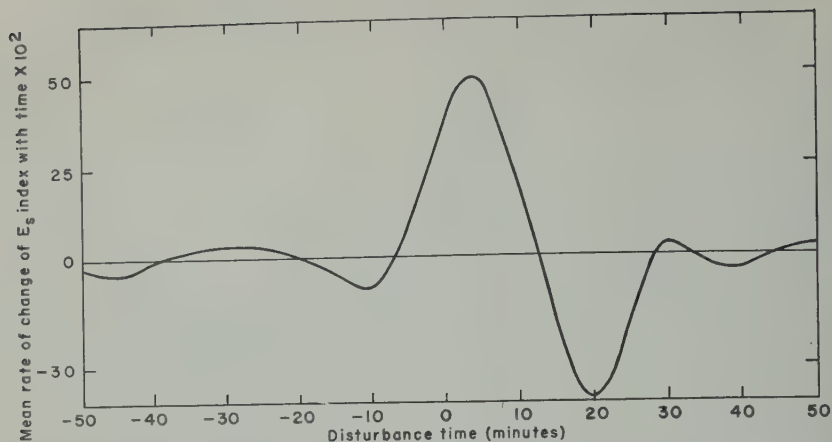


Fig. 4. Variation of rate of change of  $E_s$  index with disturbance time for all disturbances in significance  $S > 1.1 \times 10^{-2}$  (indices)<sup>2</sup> (min)<sup>-1</sup>.

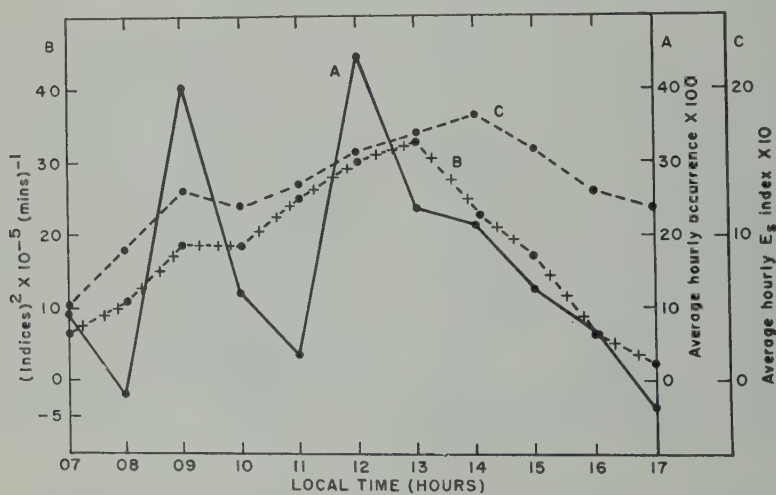


Fig. 5. A, diurnal variation of mean hourly value of  $S$ . B, diurnal variation of frequency of occurrence of TID's. C, diurnal variation of mean hourly value of  $E_s$  index.

[NOTE—Author's correction received after proof: For B and A at tops of ordinate scales read A and B respectively; letters within graphs are unchanged.]

34°03'S) on July 16, 1955. An anomaly in  $F_2$  due to a traveling disturbance is evident at 1256 hours as a cusp at high virtual height near penetration frequency. At 1258 hours the cusp complexity has fallen approximately 75 km in height and the  $E_s$  critical frequency has increased from 3.5 to 4.3 Mc/s. At 1303 and 1307 hours, as the  $F_2$  trace becomes more complex after advent of the TID,  $E_s$  extraordinary is evident. The  $E_s$  critical finally decreases again to 3.9 Mc/s at 1315 hours.

Further experimental evidence of the association between the phenomena is provided by records showing an increase in  $E_s$  at a series of stations during the passage of a traveling disturbance. This has already been described [Heisler, 1959]; Figure 7 gives a typical example in ionosonde records at Hobart (42°14'51'S, 147°51'E), Canberra (35°18'S, 149°E), Sydney (33°52'S, 151°11'E) on July 23, 1955. The times of occurrence of the  $F$ -region anomalies on the records are consistent with a



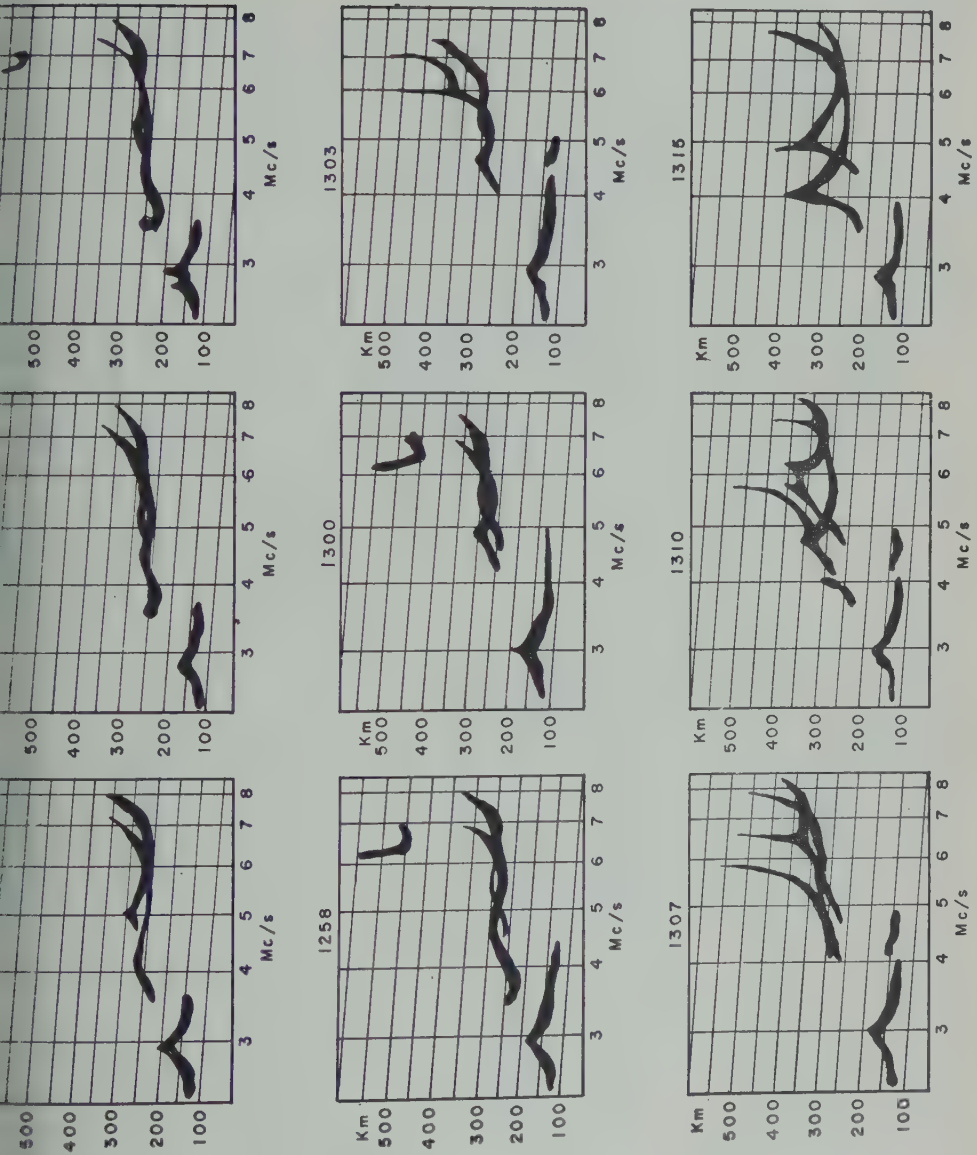
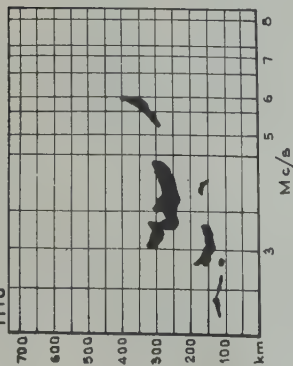
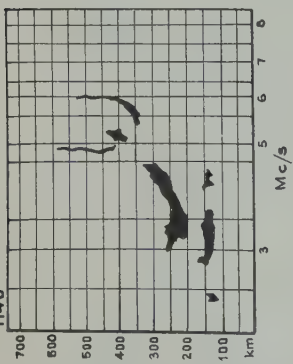


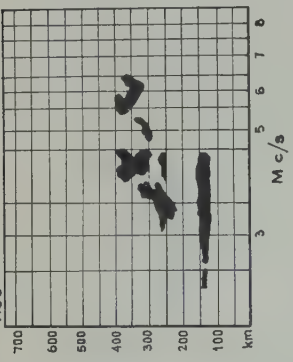
Fig. 6. Ionosonde records showing a traveling  $F_2$  disturbance at Sydney July 16, 1955, associated with an increase in sporadic-E ionization.

HOBART  
1110

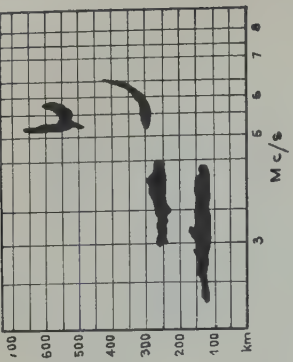
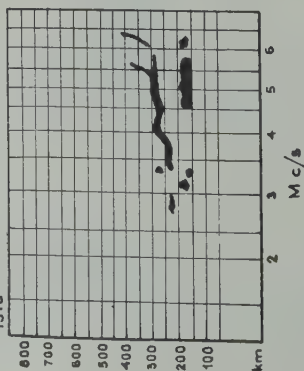
1140



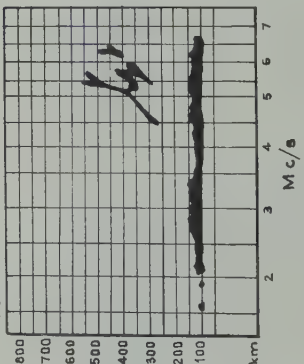
1150



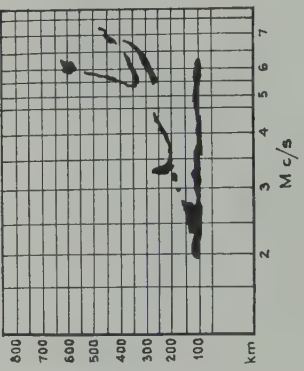
1200

CANNBERRA  
1310

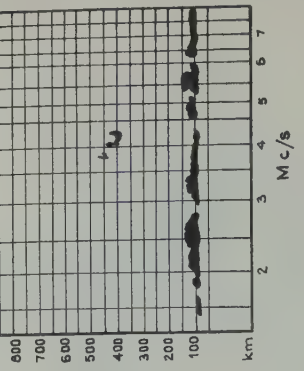
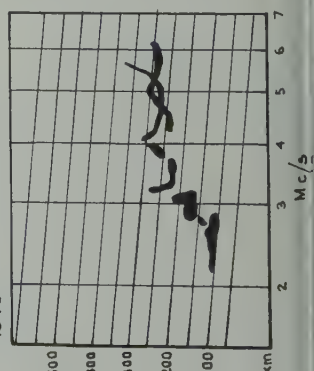
1330



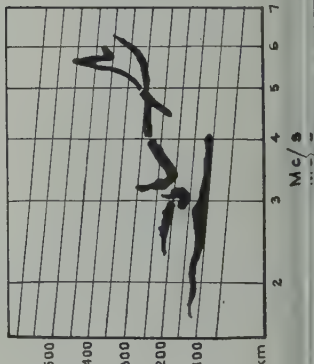
1410



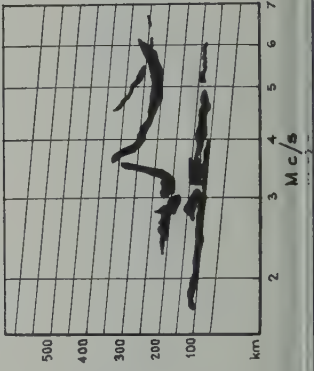
1430

SYDNEY  
1340

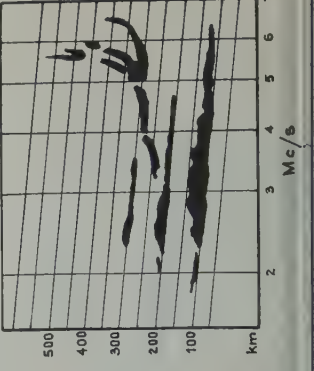
1351



1410



1426



ling due north with a speed of 7.5 km/ accompanied by the development of  $E_s$  at atations.

reful examination of individual disturb- e shows that there is always a small but rvable anomaly in  $F_1$  between disturbance and time of the related  $E_s$  effect. This d indicate that a change in ionization ay progresses from  $F_2$  maximum down to the gion and suggests that the change in  $f_oE_s$  probably due to a redistribution of ioniza-

a additional complexity in the  $F_1$  trace due ne progressive fall in height of the cusp y anomaly is often seen on records. It occurs some tes after the initial  $E_s$  anomaly and is some- s followed by another increase in  $E_s$  critical yency. This may develop like that shown e series of records in Figure 6, or it may equential in formation as originally de- ed by McNicol and Gipps [1951] and as tly observed associated with traveling-dist- ance phenomena by Bibl and Rawer [1959]. ems probable that the second maximum ob- ed in Figure 1 may be due to this even gh statistical analysis indicates that it is nificant. More detailed analysis of a er quantity of data is planned.

conclusions. A definite association between gion traveling ionospheric disturbances and as been demonstrated. The correlation is 100 per cent, and only about 1 disturbance 5 is associated with an  $E_s$  anomaly. This nences near  $F$ -disturbance time, and  $f_oE_s$  ases approximately 1 Mc/s to its maximum e 13 minutes later, returning to normal 25 tes after initiation of the  $F$  disturbance. yond suggesting that the observed  $E_s$  for-

mation may be due to a redistribution process, no attempt has been made to explain the es- tablished relationship of the two phenomena. This is being studied at present and is intended to be the basis of another paper.

*Acknowledgments.* This investigation has been carried out in the Electrical Engineering Department of the University of Sydney as part of the research program sponsored by the Radio Research Board of the Commonwealth Scientific and Industrial Research Organisation.

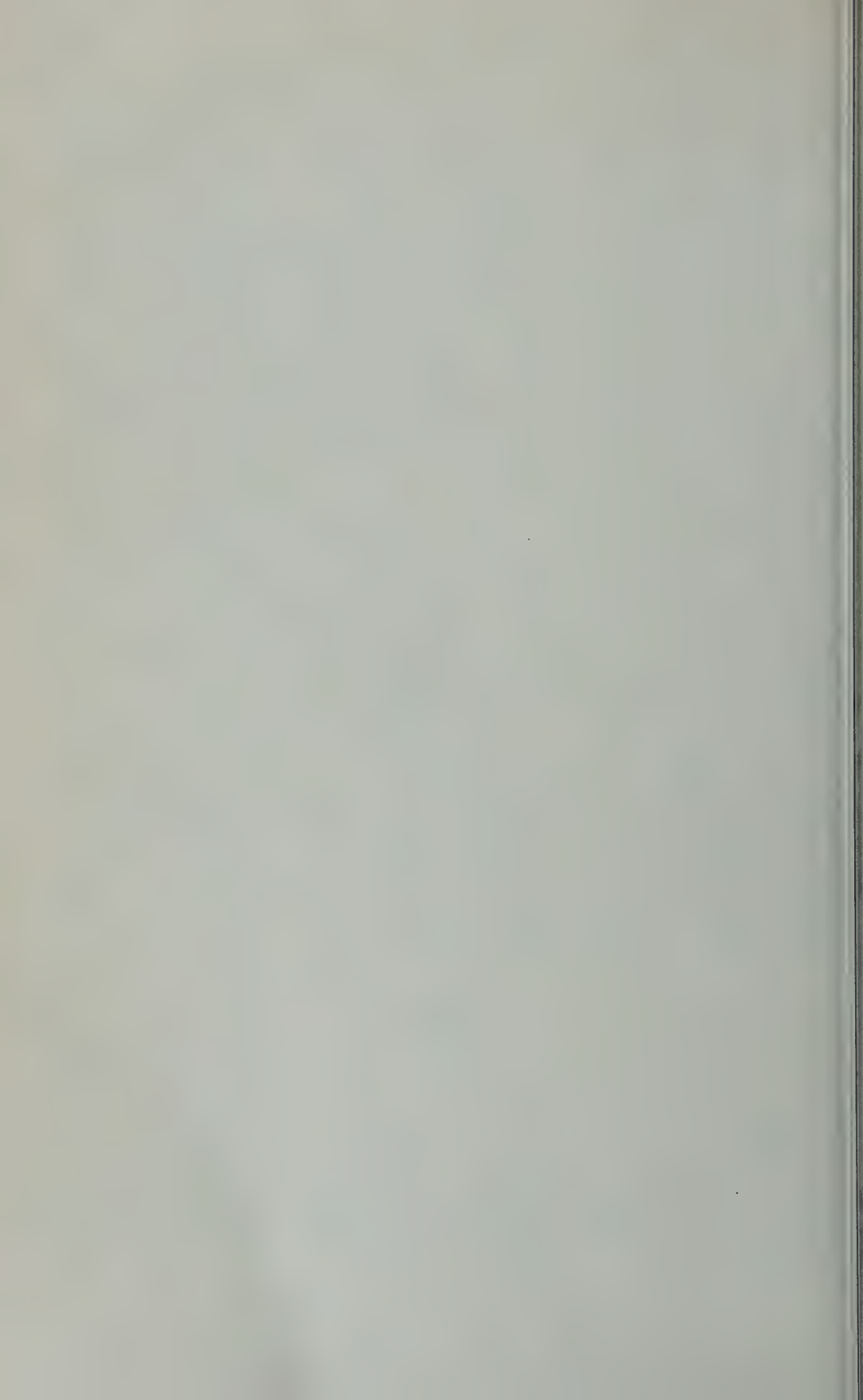
The detailed reduction of records and prepara- tion of material by the analysis staff of the Radio Research Board Laboratories, Sydney, are grate- fully acknowledged. We express our appreciation to the University of Sydney and in particular to Professor W. N. Christiansen for provision of facilities in the Electrical Engineering Department.

#### REFERENCES

- Bibl, K., and K. Rawer, Traveling disturbances originating in the outer ionosphere, *J. Geophys. Research*, **64**, 2232-2238, 1959.
- Heisler, L. H., Anomalies on ionosonde records due to travelling ionospheric disturbances, *Australian J. Phys.*, **11**, 79-90, 1958.
- Heisler, L. H., A relation between giant traveling disturbances and sporadic  $E$  ionisation, *Nature*, **184**, 1788-1789, 1959.
- McNicol, R. W. E., and G. de V. Gipps, Characteristics of the  $E_s$  region at Brisbane, *J. Geophys. Research*, **56**, 17-31, 1951.
- Munro, G. H., Travelling disturbances in the ionosphere, *Proc. Roy. Soc. London, A*, **202**, 208-223, 1950.
- Munro, G. H., and L. H. Heisler, Cusp type anomalies in variable frequency ionospheric records, *Australian J. Phys.*, **9**, 343-358, 1956.
- Thomas, J. A., Outline of suggested theories of sporadic  $E_s$ , *Agardograph* **34**, North Atlantic Treaty Organisation Advisory Group for Aeronautical Research and Development, 23-30, 1958.

(Manuscript received June 21, 1960.)





# A Variable Atmospheric-Density Model from Satellite Accelerations

LUIGI G. JACCHIA

*Smithsonian Institution Astrophysical Observatory  
Cambridge, Massachusetts*

**Abstract.** The analysis of satellite accelerations leads to an empirical formula that relates the product  $\rho H^{1/2}$  ( $\rho$  = atmospheric density,  $H$  = scale height) to the geometric height  $z$ , the 20-cm solar flux  $F_{20}$ , and the angular distance  $\psi'$  from the center of the diurnal bulge. Once the numerical parameters of this formula have been established, tables of  $\rho$  and  $H$  are computed and a separate formula is derived to represent  $\rho$  in function of the same variables.

*Solar effects in the upper atmosphere.* The acceleration of artificial satellites has revealed that the density of the upper atmosphere is due to variations in solar radiation, both wave-*Jacchia and Briggs, 1958; Priester, 1959; Jacchia, 1959b*] and corpuscular [*Jacchia, 1959c*]. Moreover, the existence of a pronounced diurnal effect has been established [*Jacchia, 1959a, d; Wyatt, 1959*]. Preliminary quantitative studies of the upper-atmospheric structure with inclusion of solar effects have been published by *Priester and Martin [1960]* and *Nicolet [1960]*.

The observed phenomena can be summarized as follows:

(a) A large diurnal effect at great heights (500 km) decreasing to almost zero at the 200-km level. The atmosphere bulges out in the radial direction of the sun, with a pronounced effect caused by the earth's rotation.

(b) Erratic fluctuations, in which the 27-day period of the solar rotation can often be recorded over many months. These fluctuations appear in a remarkable manner: the fluctuations of the solar flux at wavelengths of the order of 20 cm. At lower satellite heights (200 km) atmospheric fluctuations are relatively small and nearly independent of the position with respect to the sun. At greater heights they become quite large in the diurnal bulge but remain small in the dark hemisphere.

(c) Transient increases in the density of the upper atmosphere above 200 km during magnetic storms. Only two such events have been identified with certainty [*Jacchia, 1959c*]; in

both, the perturbation was in phase with the magnetic storm, in intensity and in duration.

The purpose of this paper is to establish an empirical set of formulas capable of representing the phenomena listed under (a) and (b) with an accuracy close to the observational error.

2. *The 20-cm solar flux.* In previous papers [*Jacchia, 1959b, 1959d*] the author had used the 10.7-cm flux, recorded daily by Dr. A. E. Covington at the National Research Council, Ottawa, for comparison with satellite accelerations. The correlation between the times of maxima and minima of the 'erratic' fluctuation in the satellite acceleration curve and of the 10.7-cm solar-flux curve is excellent. The satellite curve, however, shows also a slower fluctuation, with cycles of the order of a few months, which is not present in the curve of the 10.7-cm flux [*Jacchia, 1959d*].

Priester and Martin have used in their analysis the 20-cm flux measured daily at the Heinrich-Hertz-Institut für Schwingungsforschung in Berlin-Adlershof, which had not been available to this author. It is apparent from their diagrams that the 20-cm flux shows slow mean-level fluctuations which are just right to explain the residual oscillations in the satellite curve, in particular the dip leading to the minimum of July 1958 and the subsequent rise to a maximum in October 1958. After this discovery the author asked the Heinrich-Hertz-Institut to supply him with the 20-cm data, which he has used in the present analysis.

*Nicolet [1960]* thinks that the mean-level

fluctuations of the 20-cm solar flux must be spurious, of instrumental origin, because they do not have a counterpart in the 3.2-, 8-, 10.7-, 15-, 21-, and 30-cm fluxes measured at Berlin, Nagoya, Ottawa, and Sydney. For this reason he believes that an analysis based on the 10.7-cm flux is safer, and he distrusts Priester and Martin's findings. It must be admitted that the singularity of behavior of the 20-cm flux more than justifies Nicolet's doubts. On the other hand, we have to face the fact that the slow fluctuations of the 20-cm flux are generally reflected in the satellite accelerations: a look at Figure 1 should be convincing in this respect. If so far the present author has refrained from publishing a more detailed quantitative analysis of upper-atmosphere densities based on the 10.7-cm flux, it was mainly because of the presence of the 'erratic fluctuations of unexplained origin' [Jacchia, 1959*d*], which now seem to be just about eliminated by the use of the 20-cm flux. Without categorically excluding the possibility that some drift might be present in the 20-cm data, we are satisfied that, at least during the period covered by satellite observations, the 20-cm flux seems to have been more closely related than the 10.7-cm flux to the soft X-ray flux which is probably responsible for the atmospheric fluctuations.

The change by a factor of 2.36 in the calibration constant of the 20-cm flux starting July 1, 1958, as announced by the Heinrich-Hertz-Institut, was taken into account in our analysis. All values of the flux prior to that date were multiplied by 2.36 to bring them in alignment with the current values.

3. *Satellite accelerations.* To derive actual atmospheric densities, accelerations of the following satellites were used: 1958 $\beta_2$ , [Jacchia, 1959*d*; Briggs, 1959]; 1959 $\alpha_1$  [Jacchia and Nigam, unpublished]; 1958 $\alpha$  [Zadunaisky; Jacchia, 1959*d*]; 1958 $\delta_2$  [Y. Kozai, unpublished].

In addition, the accelerations of 1957 $\beta_1$  [Jacchia, 1958] and 1958 $\delta_1$  [Jacchia, 1959*d*] were analyzed for the oscillations only; no absolute densities were computed in view of the uncertainty in the physical characteristics of these satellites.

The accelerations of 1959 $\alpha_1$  were derived by numerical differentiation of the mean motion of the satellite, which was computed at 2-day in-

tervals by R. Nigam as part of a regular program. Since observations extending over the whole week were used for each orbit and secular acceleration was assumed to be constant in this interval, these accelerations necessarily somewhat smoothed out and times outright uncertain. The same can be said of the accelerations of 1958 $\alpha$ .

In this connection it should be pointed out that the only really homogeneous, accurately determined series of satellite accelerations in existence today is that of 1958 $\beta_2$ . These accelerations were determined by feeding the orbital elements as known functions of time into an orbital-analysis program and plotting the residuals of each observation. This method was used also for 1957 $\beta_1$ , 1958 $\delta_1$ , and 1958 $\delta_2$ , but for long periods of invisibility, the inhomogeneity of the observational material, and the sparse distribution and frequent unreliability of the orbital elements made the derivation of accelerations more uncertain for these three orbits.

4. *Densities and scale heights from accelerations.* Approximate formulas relating satellite drag to orbital elements and atmospheric parameters, derived by various authors [Sterne, 1958; Groves, 1958; King-Hele, Cook, and Walker, 1959], give essentially the same results within the limits of observational accuracy. The formula used in this paper was that of King-Hele, Cook, and Walker, namely:

$$\rho H^{1/2} = -\frac{\sqrt{2/\pi}}{3C_D} \frac{dP}{dt} \frac{m}{AF} \sqrt{\frac{e}{a}} \left[ 1 - 2ae + \frac{5e^2}{2} - \frac{H}{8ae} \left( 1 - 10e + \frac{7H}{16ae} \right) \right]$$

where

- $\rho$  = atmospheric density at satellite perigee height.
- $H$  = atmospheric scale height at satellite perigee height.
- $C_D$  = drag coefficient.
- $dP/dt$  = secular acceleration of satellite.
- $m$  = mass of satellite.
- $A$  = effective cross section of satellite.
- $F$  = factor to account for rotation of atmosphere.
- $e$  = orbital eccentricity.
- $a$  = semimajor axis of satellite orbit.

free molecular flow was assumed throughout,  $C_D$  was always taken to have the value 2. The height  $z$  corresponding to  $\rho$  was taken to be above the international geoid. In view of the paucity of the material on high-inclination satellites, no attempt was made to include latitude and seasonal effects in the analysis.

Rather than try to obtain  $\rho$  for each satellite with an assumed scale height, which would result in laborious iterations after all the material was assembled, we preferred to do all the analysis on the product  $\rho H^{1/2}$ , which can be deduced directly from the accelerations without assumptions (except for nearly negligible terms). Actual densities were derived only at the end, from the final profiles of  $\rho H^{1/2}$  in function of  $z$ . An empirical formula to describe atmospheric variations. Our formula must try to represent effects (a) and (b) of section 1. Effect (c), the corpuscular perturbation, is, of course, unpredictable; fortunately its occurrence is rare and its magnitude limited.

Effect (a) must be a function of the height  $z$ , the solar flux  $F$ , and of the angular distance from the point where the bulge is highest. It was assumed that this point is at the same altitude as the subsolar point but lags in longitude by a constant  $\lambda$ . For a point in the atmosphere whose astronomical equatorial coordinates are  $\delta_s$ , we have

$$\psi' = \sin \delta_p \sin \delta_s \\ + \cos \delta_p \cos \delta_s \cos (\alpha_p - \alpha_s - \lambda) \quad (2)$$

$\delta_s$  are the equatorial coordinates of the sun). From dynamical considerations we must assume that the lag  $\lambda$  is not necessarily the same at all heights.

An equation that will satisfactorily describe the conditions (a) and (b) can be written in the form

$$\rho H^{1/2} = f_0(z)[c_1 F_{20}^m + f_1(z)f(\psi')F_{20}] \quad (3)$$

where  $f_0(z)$ ,  $f_1(z)$ , and  $f(\psi')$  are suitable functions of  $z$  and  $\psi'$ , respectively, and  $c_1$  and  $m$  are constants.  $F_{20}$  is the 20-cm solar flux. We shall assume  $f(\psi')$  vary from 1 to zero where  $\psi'$  increases from  $0^\circ$  to  $180^\circ$ .

At  $180^\circ$  from the diurnal bulge ( $\psi' = 180^\circ$ ) the second term inside the brackets will then be zero and we can define  $f_0(z)$  as a standard

night profile of  $\rho H^{1/2}$  in function of  $z$  for a fixed value of  $F_{20}$ . The function  $f_1(z)$  is the amplitude of the diurnal effect and must be small for  $z$  close to 200 km; from satellites with perigee at this height we can thus determine the value of the exponent,  $m$ , which was introduced to fit the observed amplitude of the erratic fluctuations at low heights.

For satellite 1957 $\beta_1$  a value of  $m = 0.7$  seems to be satisfactory; for 1958 $\delta_1$  and 1958 $\delta_2$  we find that  $m = 1$  gives a sufficiently good fit. For simplicity we have assumed, provisionally,  $m = 1$ ,  $c_1 = 1$ .

A function that can describe a diurnal bulge of any degree of sharpness is

$$f(\psi') = \cos^n (\psi'/2) \quad (4)$$

The exponent  $n$  is best determined from high satellites, where the diurnal effect is large. From satellite 1958 $\beta_2$ , 1959 $\alpha_3$ , and 1958 $\alpha$  we find that  $n = 6$  gives a very satisfactory fit. Additional satellite data will probably show that  $n$  itself is a function of  $z$ .

If we express  $F_{20}$  in units of  $100 \times 10^{-22}$  watt/m<sup>2</sup>/cycle/sec and define  $f_0(z)$  as the night profile of  $\rho H^{1/2}$  for  $F_{20} = 1$ , we can write

$$\rho H^{1/2} = f_0(z)F_{20}[1 + f_1(z) \cos^6 (\psi'/2)] \quad (5)$$

Table 1 gives the values of  $f_0(z)$  and  $f_1(z)$  deduced from the various satellites, together with the best value of the lag angle  $\lambda$ . All quantities are expressed in the cgs system except the perigee heights  $z$ , which are in kilometers. For each satellite the accelerations were reduced to a standard perigee height.

From the observed values, the following expressions were found for  $f_0(z)$  and  $f_1(z)$ :

$$\log f_0(z) = -12.475 - 0.0019z \\ + 6.01 \exp (-0.0027z) \quad (6)$$

$$f_1(z) = 0.185[\exp (0.006z) - 2] \quad (7)$$

( $z$  is always expressed in kilometers;  $200 < z < 700$ ).

The values of  $f_0(z)$  and  $f_1(z)$  computed with equations 6 and 7 are to be found, for comparison, at the right-hand side of Table 1.

The final form of equation 4 becomes, then:

$$\rho H^{1/2} = f_0(z)F_{20}\{1 + 0.185 \\ \cdot [\exp (0.006z) - 2] \cos^6 (\psi'/2)\} \quad (8)$$



TABLE 1. Basic Atmospheric Data from Satellites

| Satellite       | $A/m$ | $z$ ,<br>km | Observed      |          | $\lambda$ | Computed      |          |
|-----------------|-------|-------------|---------------|----------|-----------|---------------|----------|
|                 |       |             | $\log f_0(z)$ | $f_1(z)$ |           | $\log f_0(z)$ | $f_1(z)$ |
| 1958 $\beta_2$  | 4.0   | 655         | -12.70        | 9.0      | 25°       | -12.688       | 9.00     |
| 1959 $\alpha_1$ | 4.4   | 562         | -12.22        | 4.9      | 30°       | -12.220       | 5.00     |
| 1958 $\alpha$   | 4.9   | 353         | -10.83        | 1.2      | 30°       | -10.825       | 1.10     |
| 1958 $\delta_2$ | 45.0  | 210         | -9.46         | 0.2-0.4  | ...       | -9.463        | 0.20     |

TABLE 2. Atmospheric Density Profile Computed from Equations 6 and 8

| Night ( $\psi' = 180^\circ$ ) |                               |             |              |              | Day ( $\psi' = 0^\circ$ ) |             |              |              |
|-------------------------------|-------------------------------|-------------|--------------|--------------|---------------------------|-------------|--------------|--------------|
| $z$ ,<br>km                   | $\log (\rho H^{1/2})$         |             | $\log \rho$  |              | $\log (\rho H^{1/2})$     |             | $\log \rho$  |              |
|                               | $F_{20} = 1$<br>[ $=f_0(z)$ ] | $H$ ,<br>km | $F_{20} = 1$ | $F_{20} = 3$ | $F_{20} = 1$              | $H$ ,<br>km | $F_{20} = 1$ | $F_{20} = 3$ |
|                               |                               |             |              |              |                           |             |              |              |
| 200                           | -9.353                        | 36.6        | -12.635      | -12.158      | -9.258                    | 40.7        | -12.563      | -12.00       |
| 220                           | -9.575                        | 38.3        | -12.867      | -12.389      | -9.453                    | 43.0        | -12.770      | -12.20       |
| 240                           | -9.787                        | 40.0        | -13.088      | -12.611      | -9.638                    | 45.5        | -12.966      | -12.40       |
| 260                           | -9.990                        | 41.8        | -13.301      | -12.824      | -9.811                    | 48.1        | -13.152      | -12.60       |
| 280                           | -10.185                       | 43.6        | -13.505      | -13.028      | -9.975                    | 50.9        | -13.328      | -12.80       |
| 300                           | -10.371                       | 45.4        | -13.700      | -13.223      | -10.128                   | 53.8        | -13.494      | -13.00       |
| 320                           | -10.550                       | 47.3        | -13.887      | -13.410      | -10.273                   | 57.0        | -13.651      | -13.20       |
| 340                           | -10.721                       | 49.3        | -14.067      | -13.590      | -10.409                   | 60.4        | -13.799      | -13.40       |
| 360                           | -10.885                       | 51.3        | -14.240      | -13.763      | -10.536                   | 63.9        | -13.939      | -13.60       |
| 380                           | -11.043                       | 53.4        | -14.406      | -13.929      | -10.656                   | 67.7        | -14.071      | -13.80       |
| 400                           | -11.194                       | 55.5        | -14.566      | -14.089      | -10.768                   | 71.8        | -14.195      | -14.00       |
| 420                           | -11.339                       | 57.7        | -14.720      | -14.243      | -10.872                   | 76.1        | -14.313      | -14.20       |
| 440                           | -11.479                       | 59.9        | -14.868      | -14.391      | -10.971                   | 80.7        | -14.424      | -14.40       |
| 460                           | -11.613                       | 62.2        | -15.010      | -14.533      | -11.063                   | 85.5        | -14.528      | -14.60       |
| 480                           | -11.742                       | 64.5        | -15.147      | -14.670      | -11.149                   | 90.6        | -14.627      | -14.80       |
| 500                           | -11.867                       | 66.9        | -15.280      | -14.802      | -11.229                   | 96.1        | -14.720      | -15.00       |
| 520                           | -11.987                       | 69.3        | -15.407      | -14.930      | -11.304                   | 102.0       | -14.808      | -15.20       |
| 540                           | -12.102                       | 71.8        | -15.530      | -15.053      | -11.374                   | 108.2       | -14.891      | -15.40       |
| 560                           | -12.214                       | 74.3        | -15.649      | -15.172      | -11.439                   | 114.8       | -14.968      | -15.60       |
| 580                           | -12.322                       | 76.8        | -15.764      | -15.287      | -11.500                   | 121.7       | -15.042      | -15.80       |
| 600                           | -12.426                       | 79.4        | -15.875      | -15.398      | -11.556                   | 129.0       | -15.111      | -16.00       |
| 620                           | -12.526                       | 82.0        | -15.983      | -15.506      | -11.609                   | 136.8       | -15.177      | -16.20       |
| 640                           | -12.623                       | 84.7        | -16.087      | -15.610      | -11.658                   | 145.0       | -15.239      | -16.40       |
| 660                           | -12.718                       | 87.4        | -16.188      | -15.711      | -11.703                   | 153.7       | -15.297      | -16.60       |
| 680                           | -12.809                       | 90.0        | -16.286      | -15.809      | -11.745                   | 162.9       | -15.352      | -16.80       |
| 700                           | -12.897                       | 92.7        | -16.381      | -15.904      | -11.784                   | 172.7       | -15.404      | -17.00       |

lag angle  $\lambda$  in  $\psi'$  can be assumed to be between  $25^\circ$  and  $30^\circ$  for all heights;  $f_0(z)$  is to be taken from equation 6.

Table 2 gives night and day profiles of  $\rho$  and  $\rho_0$  for  $F_{20} = 1$  and  $F_{20} = 3$ , as computed by equation 8. The  $\rho$  profile for  $\psi' = 180^\circ$  and  $F_{20} = 1$  is well represented by the equation

$$\rho_0(z) = -16.021 - 0.001985z + 6.363 \exp(-0.0026z) \quad (9)$$

expressed in kilometers;  $200 < z < 700$ . The general equation for  $\rho$  can be written in the same form as equation 8:

$$\rho_0(z)F_{20}\{1 + 0.19 \exp(0.0055z) - 1.9\} \cos^6(\psi'/2) \quad (10)$$

The numerical coefficients in equations 6, 8, 9, and 10 must, of course, be considered only provisional and susceptible of considerable improvement when more satellite acceleration material is analyzed.

Table 3 gives height profiles of the diurnal density for  $F_{20} = 2$ , i.e., the height at which a given density is reached as a function of the angular distance  $\psi'$  from the center of the bulge. These profiles are illustrated graphically in Figure 3. Figures 1 and 2 show a comparison between observed and computed accelerations for satellites 1958 $\beta_2$  and 1959 $\alpha_1$ .

*Remarks and conclusions.* The approach used in this paper is primarily descriptive; theoretical inferences on atmospheric parameters other than the density have been left out, since they cannot be derived without some degree of

speculation. Equation 10, or improved versions of it, should prove useful for a quick computation of corrections to be applied to a standard atmospheric density profile at a given time and place.

Owing to the nature of the problem, the relative accuracy of the densities derived from our formulas is considerably greater than that of the corresponding scale heights—in particular, at the two extremes of the height zone covered by satellites. A critical point occurs around 200 km, where the density profile derived from satellites should be connected with that obtained from rockets. This operation may not be very difficult for a *mean* density profile; when, however, a particular profile such as our  $\rho_0(z)$  is chosen, it becomes necessary to know how the amplitude of the erratic fluctuations related to  $F_{20}$  varies with height in the region from 100 to 200 km.

It should not be forgotten, of course, that the diurnal effect, or any other atmospheric effect dependent on geographical position, is bound to be somewhat smoothed out or even distorted in its action on satellite motions, especially if the orbital eccentricity is small (fortunately this is not true of the higher satellites investigated in this paper). This situation occurs because the formulas used to derive the atmospheric density at perigee assume that the atmosphere is spherically symmetric.

Equation 10 consists of two terms, of which the first,  $\rho_0(z)F_{20}$ , is independent of geographical position and does not affect the scale height of the atmosphere. This would imply an absorption

TABLE 3. Height Profile (km) of Surfaces of Equal Density in Function of the Angular Distance  $\psi'$  from the Center of the Diurnal Bulge ( $F_{20} = 2$ )

| $\log \rho$ | $\psi'$   |            |            |            |            |            |            |             |             |             |
|-------------|-----------|------------|------------|------------|------------|------------|------------|-------------|-------------|-------------|
|             | $0^\circ$ | $15^\circ$ | $30^\circ$ | $45^\circ$ | $60^\circ$ | $75^\circ$ | $90^\circ$ | $120^\circ$ | $150^\circ$ | $180^\circ$ |
| -15.5       | 938.9     | 919.6      | 868.9      | 802.1      | 733.3      | 673.5      | 630.1      | 592.0       | 586.7       | 586.6       |
| -15.0       | 662.2     | 655.2      | 635.7      | 607.4      | 575.9      | 547.1      | 525.6      | 506.1       | 503.4       | 503.3       |
| -14.5       | 518.0     | 514.2      | 503.5      | 487.9      | 470.6      | 454.8      | 443.1      | 432.4       | 430.9       | 430.9       |
| -14.0       | 418.0     | ...        | 409.3      | ...        | 389.9      | ...        | 374.1      | 368.2       | 367.4       | 367.2       |
| -13.5       | 340.8     | ...        | 335.5      | ...        | 323.9      | ...        | 314.7      | 311.2       | 310.8       | 310.7       |
| -13.0       | 277.7     | ...        | 274.5      | ...        | 267.6      | ...        | 262.3      | 260.3       | 260.0       | 260.0       |
| -12.5       | 224.0     | ...        | 222.2      | ...        | 218.4      | ...        | 215.5      | 214.4       | 214.3       | 214.3       |

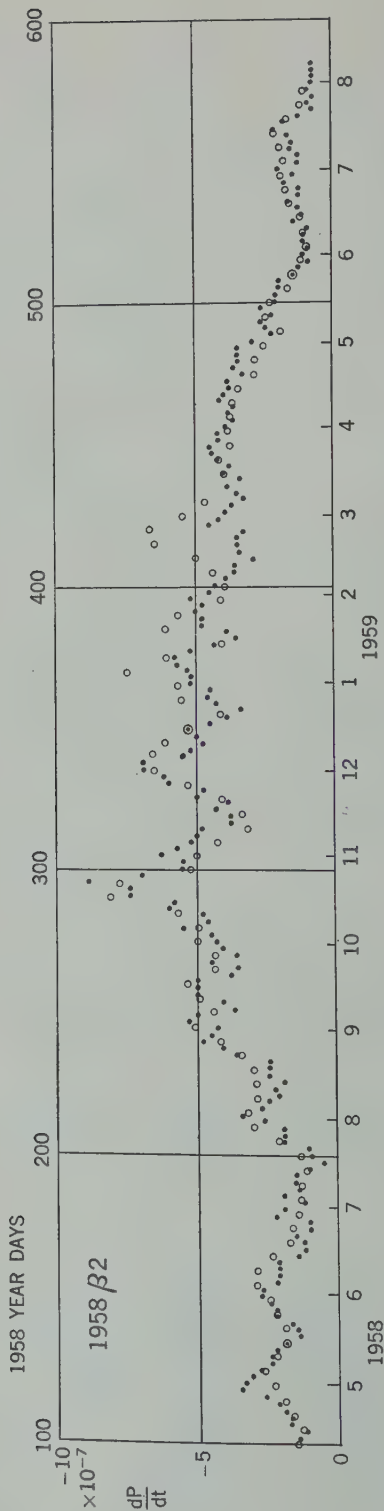


Fig. 1. Observed accelerations of satellite 1958 $\beta_2$  (dots) compared with accelerations computed by equation 8 (open circles), with lag angle  $\lambda = 25^\circ$ .

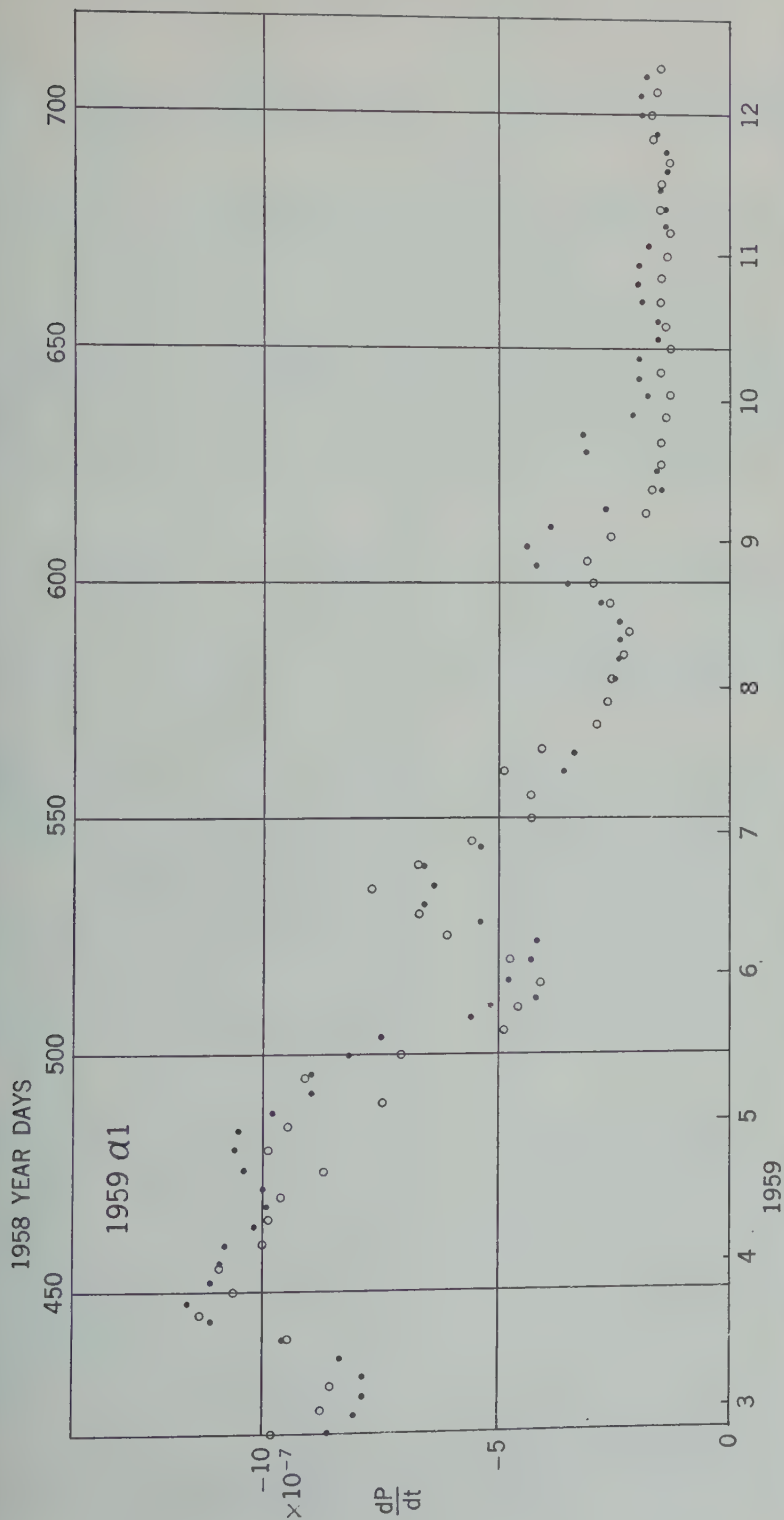


Fig. 2. Observed accelerations of satellite 1959 $\alpha_1$  (dots) compared with accelerations computed by equation 8 (open circles), with lag angle  $\lambda = 30^\circ$ . No 20-cm flux data were available after Aug. 31, 1959 ( $=$  1958 year day 608); after that date the 10.7-cm flux from Ottawa was used, multiplied by an empirical factor 0.85 to adjust its mean level to that of the 20-cm flux.



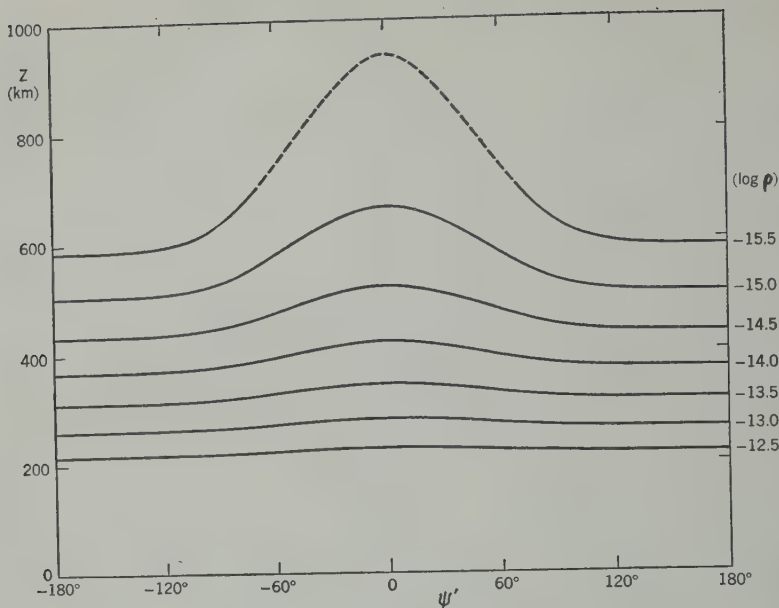


Fig. 3. Heights of surfaces of equal density above a great circle across the diurnal bulge, computed for a 20-cm solar flux of  $200 \times 10^{-22}$  watt/m<sup>2</sup> cycle/sec. The value of  $\log \rho$  corresponding to each curve is marked on the right-hand margin.

of the effective solar radiation that takes place entirely below the 200-km level and a uniform distribution of the absorbed energy throughout both the bright and the dark hemispheres. The second term of equation 10 vanishes at  $z = 117$  km, when  $\exp(0.0055z) = 1.9$ . The location of this vanishing point is highly uncertain, but it is tempting to identify this height with the mean height of the absorption zone. The diurnal effect can easily be explained by the mechanism proposed by Nicolet [1960].

#### REFERENCES

- Briggs, R. E., A table of times of perigee passage for satellite 1958 $\beta_2$ , *Smithsonian Astrophys. Observatory Spec. Rept.* 30, 1959.
- Groves, G. V., Effect of the earth's equatorial bulge on the lifetime of artificial satellites and its use in determining atmosphere scale heights, *Nature*, 181, 1055, 1958.
- Jacchia, L. G., Orbital results for satellite 1957 $\beta_1$ , *Smithsonian Astrophys. Observatory Spec. Rept.* 13, 1958.
- Jacchia, L. G., The diurnal effect in the orbital acceleration of satellite 1957 $\beta_1$ , *Smithsonian Astrophys. Observatory Spec. Rept.* 20, 1959a.
- Jacchia, L. G., Two atmospheric effects in the orbital acceleration of artificial satellites, *Nature*, 183, 526-527, 1959b.
- Jacchia, L. G., Corpuscular radiation and the deceleration of artificial satellites, *Nature*, 166, 1662-1663, 1959c.
- Jacchia, L. G., Solar effects on the acceleration of artificial satellites, *Smithsonian Astrophys. Observatory Spec. Rept.* 29, 1959d.
- Jacchia, L. G., and R. E. Briggs, Orbital acceleration of satellite 1958 $\beta_2$ , *Smithsonian Astrophys. Observatory Spec. Rept.* 18, 1958.
- King-Hele, D. G., G. E. Cook, and D. M. Walker, Contraction of satellite orbits under the influence of air drag, part 1, *Roy. Air Force Establishment (Farnborough) Tech. Note* 533, 1959.
- Nicolet, M., Les variations de la densité et le transport de chaleur par conduction dans l'atmosphère supérieure, *Centre natl. recherches espace Bruxelles note prélim.* 5, 1960.
- Priester, W., Sonnenaktivität und Abbremsung von Erdsatelliten, *Naturwissenschaften*, 46, 197, 1959.
- Priester, W., and H. A. Martin, Solare und zeitliche Effekte in der Hochatmosphäre, *Beobachtungen an künstlichen Satelliten*, *Univ.-Sternwarte Bonn*, 29, in press, 1960.
- Sterne, T. E., Formula for inferring atmospheric density from the motion of artificial earth satellites, *Science*, 127, 1245, 1958.
- Wyatt, S. P., Solar effects in the motion of the Vanguard, *Nature*, 184, Suppl., 351, 1959.

(Manuscript received May 20, 1960.)

# On the Motion of a Satellite in an Asymmetrical Gravitational Field

PETER MUSEN

*Theoretical Division, Goddard Space Flight Center  
National Aeronautics and Space Administration  
Washington, D. C.*

**Abstract.** Existing satellite data have a tracking precision and a limited interval of observation sufficient only for the determination of the zonal harmonics in the geoid. With the availability of more accurate data, extended over longer intervals of time, it will also become possible to determine the tesseral harmonics in the gravitational field. The present theory, developed in anticipation of the availability of these data, extends the work published by the author recently to permit the easy inclusion of the influence of tesseral harmonics, and consequently also of gravitational anomalies, in the motion of an artificial satellite.

**Introduction.** This article represents a generalization of the author's earlier theory of artificial satellites [Musen, 1959]. In the previous theory the development of the earth's potential contained only the zonal harmonics. At present the zonal harmonics are the only ones that can be derived from observation; however, it can be expected that, with the launching of geodetic satellites and the availability of more precise observations extended over longer intervals of time, it will become possible to establish the existence of tesseral harmonics and to obtain a more complete description of the earth's potential. The general theory presented in this article is developed in anticipation of the necessity of correcting the present models of the geoid and also for the correction of the existing marine program. This theory, like the previous one, is devised for the use of large-scale electronic computers, but the assumption on the gravitational potential is more general, namely, that the central body has an irregular shape and an irregular distribution of mass.

In addition, we suppose that the distance of the satellite from the body is such that the rotating Hansen ellipse can be considered as a starting point for the method of iteration. This ellipse lies in the osculating orbit plane, has a constant shape, and is rotating uniformly with respect to the disturbed eccentric anomaly of the satellite.

**Development of the disturbing function.** The relations used in this exposition are the same as in the previous article [Musen, 1959], but

their meaning is sometimes slightly different. The basic variable will be the disturbed eccentric anomaly  $E$  of the satellite. As before, in order to separate the elliptic motion from the perturbations, an auxiliary satellite, moving on the rotating ellipse, will be introduced [Hansen, 1838] as a temporary substitute for the real satellite.

The 'elliptical' eccentric anomaly  $F$  is used to obtain the components of the disturbing force and in the integrating of the basic  $W$  function. After performing these operations  $F$  is replaced by  $E$  again. The 'non-elliptical' eccentric anomaly  $E$  plays the role equivalent to time in the development of the perturbations. The disturbing function will be introduced, not in its classical form, but in the modified form with the separated eccentric anomalies.

Let  $\rho$ ,  $\lambda$ , and  $\phi$  be the spherical coordinates of the *auxiliary* satellite with respect to the system  $X'Y'Z'$  rigidly connected with the body. Let  $1 + \nu$ , as usual, be the perturbations of the radius vector of the *real* satellite. The form of the disturbing function which will be used at the outset is

$$\Omega = \sum_{n=1}^{\infty} \sum_{m=0}^n \frac{\cos^m \phi}{(1 + \nu)^{n+1}} \cdot \frac{1}{\rho^{n+1}} \cdot \frac{d^m P_n(\psi)}{d\psi^m} \cdot (A_{nm} \cos m\lambda + B_{nm} \sin m\lambda) \quad (1)$$

Let  $(XY)$  be the system of coordinates rigidly connected with the osculating orbit plane. Designating, as before, the mean motion of perigee in the osculating orbit plane by  $y$ , we have

the standard system of equations, giving the polar coordinates  $\rho$  and  $v$  of the auxiliary satellite with respect to  $XY$

$$\rho/a_0 = 1 - e_0 \cos F$$

$$v = \varphi + \pi_0 + \gamma \Delta E, \quad \Delta E = E - E_0 \quad (2)$$

$$(\rho/a_0) \cos \varphi = \cos F - e_0$$

$$(\rho/a_0) \sin \varphi = \sqrt{1 - e_0^2} \sin F \quad (2')$$

The position of  $(XY)$  with respect to  $(X'Y'Z')$  is given by the standard Eulerian angles:

- $i$ , the inclination;
- $\theta$ , the longitude of the ascending node;
- $\sigma$ , the position of the departure point.

The position of  $(X'Y'Z')$  with respect to the inertial system  $(xyz)$  is given by three similar Eulerian angles:  $\vartheta, z, \zeta$ .

In this more general theory we define two quantities  $N$  and  $K$  and the mean elements in a slightly different way from before, because the motion of the system  $(X'Y'Z')$  must be taken into consideration.

We put

$$2N = \sigma_0 + \theta_0 - \sigma - \theta - (2\alpha + m) \cdot \Delta E \quad (3)$$

$$2K = \sigma_0 - \theta_0 - \sigma + \theta + (2\eta + m) \cdot \Delta E$$

$2\alpha + m$  and  $2\eta + m$  are determined in such a way that  $N$  and  $K$  do not contain any secular terms.

The values of the mean elements  $(\sigma)$ ,  $(\theta)$ ,  $(\omega)$  are

$$(\sigma) = \sigma_0 - (\alpha - \eta) \cdot \Delta E$$

$$(\theta) = \theta_0 - (\alpha + \eta + m) \cdot \Delta E \quad (4)$$

$$(\omega) = (\pi_0 - \sigma_0) + (\gamma + \alpha - \eta) \cdot \Delta E$$

The remaining mean elements are  $a_0, e_0, i_0$ . They are constants of the integration. We deduce

$$\sigma = (\sigma) - (N + K)$$

$$\theta = (\theta) - (N - K) \quad (5)$$

$$v - \sigma = \varphi + (\omega) + N + K$$

$$v - \sigma + \theta = \varphi + (\omega) + (\theta) + 2K \quad (6)$$

$$v - \sigma - \theta = \varphi + (\omega) - (\theta) + 2N$$

The equations

$$\cos \phi \cos \lambda = \cos (v - \sigma) \cos \theta$$

$$- \sin (v - \sigma) \sin \theta \cos i$$

$$\cos \phi \sin \lambda = \cos (v - \sigma) \sin \theta$$

$$+ \sin (v - \sigma) \cos \theta \cos i$$

$$\psi = \sin \phi = \sin i \sin (v - \sigma)$$

can be written in the form

$$\cos \phi \cos \lambda = \cos^2 (i/2) \cos (v - \sigma + \theta)$$

$$+ \sin^2 (i/2) \cos (v - \sigma - \theta)$$

$$\cos \phi \sin \lambda = \cos^2 (i/2) \sin (v - \sigma + \theta)$$

$$- \sin^2 (i/2) \sin (v - \sigma - \theta)$$

$$\psi = 2 \sin (i/2) \cos (i/2) \sin (v - \sigma)$$

Introducing the parameters

$$\lambda_1 = \sin (i/2) \cos N, \lambda_3 = \cos (i/2) \sin N$$

$$\lambda_2 = \sin (i/2) \sin N, \lambda_4 = \cos (i/2) \cos N$$

and taking equations 6 and 8 into consideration we obtain

$$(\rho/a_0) \cos \phi \cos \lambda$$

$$= (\lambda_4^2 - \lambda_3^2) \cdot (\rho/a_0) \cos [\varphi + (\omega) + (\theta)]$$

$$- 2\lambda_3\lambda_4 \cdot (\rho/a_0) \sin [\varphi + (\omega) + (\theta)]$$

$$+ (\lambda_1^2 - \lambda_2^2) \cdot (\rho/a_0) \cos [\varphi + (\omega) - (\theta)]$$

$$- 2\lambda_1\lambda_2 \cdot (\rho/a_0) \sin [\varphi + (\omega) - (\theta)]$$

$$(\rho/a_0) \cos \phi \sin \lambda$$

$$= (\lambda_4^2 - \lambda_3^2) \cdot (\rho/a_0) \sin [\varphi + (\omega) + (\theta)]$$

$$+ 2\lambda_3\lambda_4 \cdot (\rho/a_0) \cos [\varphi + (\omega) + (\theta)]$$

$$- (\lambda_1^2 - \lambda_2^2) \cdot (\rho/a_0) \sin [\varphi + (\omega) - (\theta)]$$

$$- 2\lambda_1\lambda_2 \cdot (\rho/a_0) \cos [\varphi + (\omega) - (\theta)]$$

$$\psi = 2(a_0/\rho) \{ (\lambda_1\lambda_4 - \lambda_2\lambda_3)$$

$$\cdot (\rho/a_0) \sin [\varphi + (\omega)]$$

$$+ (\lambda_2\lambda_4 + \lambda_1\lambda_3)$$

$$\cdot (\rho/a_0) \cos [\varphi + (\omega)] \}$$

making the formulas

$$\cos(\varphi + \beta) = \cos^2(\varphi_0/2) \cos(F + \beta) + \sin^2(\varphi_0/2) \cos(F - \beta) - \sin \varphi_0 \cos \beta \quad (13)$$

$$\sin(\varphi + \beta) = \cos^2(\varphi_0/2) \sin(F + \beta) - \sin^2(\varphi_0/2) \sin(F - \beta) - \sin \varphi_0 \sin \beta \quad (14)$$

consideration, we have

| $\gamma = iF + j(\omega) + k(\theta)$ |     |     |   | Coefficients of<br>$\cos \gamma$                  | Coefficients of<br>$\sin \gamma$ |
|---------------------------------------|-----|-----|---|---|----------------------------------|
| $i$                                   | $j$ | $k$ |   |   |                                  |
| +1                                    | +1  | +1  | $+(\lambda_4^2 - \lambda_3^2) \cos^2 \frac{\varphi_0}{2}$ | $-2\lambda_3\lambda_4 \cos^2 \frac{\varphi_0}{2}$ |                                  |
| +1                                    | -1  | -1  | $+(\lambda_4^2 - \lambda_3^2) \sin^2 \frac{\varphi_0}{2}$ | $+2\lambda_3\lambda_4 \sin^2 \frac{\varphi_0}{2}$ |                                  |
| 0                                     | +1  | +1  | $-(\lambda_4^2 - \lambda_3^2) \sin \varphi_0$             | $+2\lambda_3\lambda_4 \sin \varphi_0$             | (15)                             |
| +1                                    | -1  | +1  | $+(\lambda_1^2 - \lambda_2^2) \sin^2 \frac{\varphi_0}{2}$ | $+2\lambda_1\lambda_2 \sin^2 \frac{\varphi_0}{2}$ |                                  |
| +1                                    | +1  | -1  | $+(\lambda_1^2 - \lambda_2^2) \cos^2 \frac{\varphi_0}{2}$ | $-2\lambda_1\lambda_2 \cos^2 \frac{\varphi_0}{2}$ |                                  |
| 0                                     | +1  | -1  | $-(\lambda_1^2 - \lambda_2^2) \sin \varphi_0$             | $+2\lambda_1\lambda_2 \sin \varphi_0$             |                                  |

| $\gamma = iF + j(\omega) + k(\theta)$ |     |     |   | Coefficients of<br>$\cos \gamma$                          | Coefficients of<br>$\sin \gamma$ |
|---------------------------------------|-----|-----|---|---|----------------------------------|
| $i$                                   | $j$ | $k$ |   |   |                                  |
| +1                                    | +1  | +1  | $+2\lambda_3\lambda_4 \cos^2 \frac{\varphi_0}{2}$ | $+(\lambda_4^2 - \lambda_3^2) \cos^2 \frac{\varphi_0}{2}$ |                                  |
| +1                                    | -1  | -1  | $+2\lambda_3\lambda_4 \sin^2 \frac{\varphi_0}{2}$ | $-(\lambda_4^2 - \lambda_3^2) \sin^2 \frac{\varphi_0}{2}$ |                                  |
| 0                                     | +1  | +1  | $-2\lambda_3\lambda_4 \sin \varphi_0$             | $-(\lambda_4^2 - \lambda_3^2) \sin \varphi_0$             | (16)                             |
| +1                                    | +1  | -1  | $-2\lambda_1\lambda_2 \cos^2 \frac{\varphi_0}{2}$ | $-(\lambda_1^2 - \lambda_2^2) \cos^2 \frac{\varphi_0}{2}$ |                                  |
| +1                                    | -1  | +1  | $-2\lambda_1\lambda_2 \sin^2 \frac{\varphi_0}{2}$ | $+(\lambda_1^2 - \lambda_2^2) \sin^2 \frac{\varphi_0}{2}$ |                                  |
| 0                                     | +1  | -1  | $+2\lambda_1\lambda_2 \sin \varphi_0$             | $+(\lambda_1^2 - \lambda_2^2) \sin \varphi_0$             |                                  |

The coefficients of  $\sin \gamma$  or  $\cos \gamma$  in these tables are themselves the series in the same arguments  $F$ ,  $(\omega)$ , and  $(\theta)$ , and also in the arguments introduced by the motion of the instantaneous axis of rotation.

| $iF + j(\omega) + k(\theta)$ |     |     |   | $\cos \gamma$   | $\cos \gamma$ |
|------------------------------|-----|-----|---|---|---------------|
| $i$                          | $j$ | $k$ |   |   |               |
| +1                           | +1  | 0   | $+(\lambda_2\lambda_4 + \lambda_1\lambda_3) \cos^2 \frac{\varphi_0}{2}$ | $+(\lambda_1\lambda_4 - \lambda_2\lambda_3) \cos^2 \frac{\varphi_0}{2}$ |               |
| +1                           | -1  | 0   | $+(\lambda_2\lambda_4 + \lambda_1\lambda_3) \sin^2 \frac{\varphi_0}{2}$ | $-(\lambda_1\lambda_4 - \lambda_2\lambda_3) \sin^2 \frac{\varphi_0}{2}$ |               |
| 0                            | +1  | 0   | $-(\lambda_2\lambda_4 + \lambda_1\lambda_3) \sin \varphi_0$             | $-(\lambda_1\lambda_4 - \lambda_2\lambda_3) \sin \varphi_0$             | (17)          |



and

$$\frac{a_0}{\rho} = \frac{2}{\sqrt{1-e_0^2}} \left( \frac{1}{2} + \operatorname{tg} \frac{\varphi_0}{2} \cos F + \operatorname{tg}^2 \frac{\varphi_0}{2} \cos 2F + \dots \right) \quad (18)$$

Let us put

$$X_m = \left( \frac{\rho}{a_0} \right)^m \cos^m \phi \cos m\lambda \quad (19)$$

$$Y_m = \left( \frac{\rho}{a_0} \right)^m \cos^m \phi \sin m\lambda \quad (20)$$

$$X_m + jY_m = \left( \frac{\rho}{a_0} \right)^m \cos^m \phi \exp(jm\lambda) \quad (21)$$

then we can have different systems of formulas for the computation of  $X_m$  and  $Y_m$ . These formulas are the analogues of trigonometric identities.

We have

$$X_m = X_1^m F\left(-\frac{m}{2}, \frac{1-m}{2}, \frac{1}{2}, -\frac{Y_1^2}{X_1^2}\right) \quad (22)$$

$$Y_m = m Y_1 X_1^{m-1} F\left(\frac{1-m}{2}, \frac{2-m}{2}, \frac{3}{2}, -\frac{Y_1^2}{X_1^2}\right) \quad (23)$$

or more convenient formulas are

$$X_{m+1} = X_1 X_m - Y_1 Y_m \quad (24)$$

$$Y_{m+1} = X_1 Y_m + Y_1 X_m \quad (25)$$

or

$$X_{m+1} = 2X_1 X_m - (X_1^2 + Y_1^2) X_{m-1} \quad (26)$$

$$Y_{m+1} = 2X_1 Y_m - (X_1^2 + Y_1^2) Y_{m-1} \quad (27)$$

The derivatives of the Legendre polynomials are computed by means of the Horner scheme, using formulas 17 and

$$\begin{aligned} \frac{d^m P_n(\psi)}{d\psi^m} &= \frac{1 + (-1)^{n-m}}{2} \\ &\cdot F\left(\frac{m-n}{2}, \frac{m+n+1}{2}, \frac{1}{2}, \psi^2\right) \\ &+ \frac{1 + (-1)^{n-m+1}}{2} F\left(\frac{m-n+1}{2}, \right. \\ &\left. \frac{m+n+2}{2}, \frac{3}{2}, \psi^2\right) \end{aligned} \quad (27')$$

All Fourier series introduced here are with numerical coefficients, and their multiplication is performed on electronic machine automatically, using either the program developed at the Vanguard Computing Center or the one developed by A. Smith and J. Monasterski at NASA Theoretical Division. As a result we will obtain the modified disturbing function in the form

$$\Omega = \sum_{n=1}^{\infty} \sum_{m=0}^n \left( \frac{a_0}{\rho} \right)^{m+n+1} (1+\nu)^{-n-1} \frac{d^m P_n(\psi)}{d\psi^m} \cdot (\alpha_{nm} X_m + \beta_{nm} Y_m)$$

Using the computed values of

$$\left( \frac{a_0}{\rho} \right)^{m+n+1}, \frac{d^m P_n(\psi)}{d\psi^m}, X_m, Y_m$$

which we store in the machine, we also compute

$$\begin{aligned} \rho \frac{\partial \Omega}{\partial \rho} &= - \sum_{n=1}^{\infty} \sum_{m=0}^n (n+1) \left( \frac{a_0}{\rho} \right)^{m+n+1} \\ &\cdot (1+\nu)^{-n-1} \frac{d^m P_n(\psi)}{d\psi^m} \\ &\cdot (\alpha_{nm} X_m + \beta_{nm} Y_m) \end{aligned}$$

which, evidently, has the same analytical form as equation 28. The significance of the introduction of the modified disturbing function with three parameters lies in the fact that only three angles

$E$

$$(\omega) = \omega_0 + (y + \alpha - \eta) \cdot \Delta E$$

$$\theta = \theta_0 - (\alpha + \eta + m) \cdot \Delta E$$

connected with the motion of the satellite appear in the arguments in the development of perturbations at each step of iteration, and the basic angles are linear in  $E$ . An iteration will change the mean motions of  $(\omega)$  and  $(\theta)$  slightly, but the linearity of the arguments with respect to  $E$  will never be disturbed. If we have reason to believe that no interaction of a given harmonic with other harmonics takes place, then we can simply put  $\omega = (\omega)$ ,  $\theta = (\theta)$ , and

$$\lambda_1 = \sin(i_0/2)$$

$$\lambda_2 = 0$$

$$\lambda_3 = 0$$

$$\lambda_4 = \cos(i_0/2)$$

development of this harmonic or use a simplified form of equations 15 to 17. We can, if we wish, treat this harmonic separately.

*Determination of the W function.* The basic function is obtained by means of an equation similar to equation 26 of the first article,

$$= N\Lambda\rho \frac{\partial a_0\Omega}{\partial \rho} + M\Lambda \frac{\partial a_0\Omega}{\partial F} + \frac{Sy}{\sqrt{1-e_0^2}} \quad (30)$$

as before,

$$\begin{aligned} & e_0^2 M \\ & \frac{h_0^2}{2} [-2 + 2e_0 \cos E + 2 \cos(F - E) \\ & e_0 \cos(F - 2E) - e_0 \cos F] \\ & \frac{1}{1+\nu} [2e_0^2 - 2e_0 \cos E \\ & e_0^2 \cos(F + E) \\ & (2 - e_0^2) \cos(F - E) - 2e_0 \cos F] \\ & (-1 - \frac{1}{2}e_0^2 + 2e_0 \cos E \\ & \frac{1}{2}e_0^2 \cos 2E) \end{aligned} \quad (31)$$

$$\begin{aligned} & e_0^2 N = \frac{h^2}{h_0^2} [+2e_0 \sin E \\ & - e_0 \sin F + e_0 \sin(F - 2E)] \\ & + \frac{1}{1+\nu} [-2e_0 \sin E \\ & - (2 - e_0^2) \sin(F - E) \\ & + e_0^2 \sin(F + E)] \\ & + (e_0 \sin E - \frac{1}{2}e_0^2 \sin 2E) \end{aligned} \quad (32)$$

$$= \frac{1-\nu^2}{1+W} \left( 1 + \frac{y}{\sqrt{1-e_0^2}} \cdot \frac{\bar{\rho}}{a_0} \right) \quad (33)$$

$$\frac{\rho}{a_0} \frac{\partial W}{\partial F} - \left( W + 1 + \frac{h_0}{h} \right) e_0 \sin F \quad (34)$$

and  $\bar{\varphi}$  designates the replacement operation

$$\bar{\Phi} = \varphi|_{F-E}$$

The perturbations of the radius vector  $1 + \nu$  and of the mean anomaly  $n_0 \delta z$  are also determined in the manner described in the first article.

*Perturbations of the orbit plane.* The equations for the perturbations of the orbit plane for the general case are different from the equations for the special case as given in *Musen* [1959].

The absolute angular velocity of rotation of the orbit plane is

$$hZr \quad (35)$$

where  $Z$  is the component of the disturbing force in the direction normal to the orbit plane, and

$$h = a^{-1/2}(1 - e^2)^{-1/2}$$

where  $a$  and  $e$  are the osculating elements, and  $r$  is the position vector of the satellite. Let us designate the unit vector along  $r$  by  $r^0$ , the unit vector along the line of nodes (along the intersection of the  $X'Y'$  plane with the orbit plane) by  $u^0$ .

The system of basic unit vectors connected with the system  $X'Y'Z'$  will be designated by  $i', j', k'$ , and a similar system of unit vectors connected with  $XYZ$  (the system rigidly tied to the osculating orbit plane) will be designated by  $P, Q, R$ . The angular velocity of rotation of the orbit plane relative to  $X'Y'Z'$  is

$$u^0 \frac{di}{dt} + k' \frac{d\theta}{dt} - R \frac{d\sigma}{dt} \quad (36)$$

The absolute angular velocity of rotation of the system  $X'Y'Z'$  can be represented in the form

$$(\omega_1 i' + \omega_2 j' + \omega_3 k')(dE/dt)$$

It follows from (35) to (37) that

$$\begin{aligned} u^0 \frac{di}{dE} + k' \frac{d\theta}{dE} - R \frac{d\sigma}{dE} &= hZr \frac{dt}{dE} \\ &- (\omega_1 i' + \omega_2 j' + \omega_3 k') \end{aligned} \quad (37)$$

Taking the relations

$$\begin{aligned} k' \cdot u^0 &= 0 & r^0 \cdot k' &= \sin i \sin(\nu - \sigma) \\ k \cdot u^0 &= 0 & r^0 \cdot R &= 0 \end{aligned}$$

$$\mathbf{r}' \cdot \mathbf{u}^0 = \cos(\nu - \sigma) \mathbf{i}' \cdot \mathbf{R} = \sin i \sin \theta \quad (37')$$

$$\mathbf{i}' \cdot \mathbf{u}^0 = \cos \theta \quad \mathbf{j}' \cdot \mathbf{R} = -\sin i \cos \theta$$

$$\mathbf{j}' \cdot \mathbf{u}^0 = \sin \theta \quad \mathbf{R} \cdot \mathbf{k}' = \cos i$$

into consideration, we deduce from equation 37 by forming 'dot' products with  $\mathbf{u}^0$ ,  $\mathbf{R}$ , and  $\mathbf{k}'$

$$\begin{aligned} \frac{di}{dE} &= hZr \cos(\nu - \sigma) \frac{dt}{dE} \\ &\quad - (\omega_1 \cos \theta + \omega_2 \sin \theta) \end{aligned} \quad (38)$$

$$\begin{aligned} \frac{d\theta}{dE} - \frac{d\sigma}{dE} \cos i \\ = hZr \sin(\nu - \sigma) \sin i \frac{dt}{dE} - \omega_3 \end{aligned} \quad (39)$$

$$\begin{aligned} \cos i \frac{d\theta}{dE} - \frac{d\sigma}{dE} &= -(\omega_1 \sin \theta \sin i \\ &\quad - \omega_2 \sin i \cos \theta + \omega_3 \cos i) \end{aligned} \quad (40)$$

It follows from equations 39 and 40

$$\begin{aligned} 2 \cos^2 \frac{i}{2} \frac{d}{dE} (\theta - \sigma) \\ = hZr \sin(\nu - \sigma) \sin i \frac{dt}{dE} \\ - 2 \cos^2 \frac{i}{2} \cdot \omega_3 \\ - \omega_1 \sin i \sin \theta + \omega_2 \sin i \cos \theta \end{aligned} \quad (41)$$

$$\begin{aligned} 2 \sin^2 \frac{i}{2} \frac{d}{dE} (\theta + \sigma) \\ = hZr \sin(\nu - \sigma) \sin i \frac{dt}{dE} \\ - 2 \omega_3 \sin^2 \frac{i}{2} + \omega_1 \sin i \sin \theta \\ - \omega_2 \sin i \cos \theta \end{aligned} \quad (42)$$

$$\begin{aligned} \frac{d}{dE} (\theta - \sigma) &= hZr \sin(\nu - \sigma) \operatorname{tg} \frac{i}{2} \frac{dt}{dE} - \omega_3 \\ &\quad - \omega_1 \sin \theta \cdot \operatorname{tg} \frac{i}{2} + \omega_2 \cos \theta \cdot \operatorname{tg} \frac{i}{2} \end{aligned} \quad (43)$$

$$\begin{aligned} \frac{d}{dE} (\theta + \sigma) \\ = hZr \sin(\nu - \sigma) \operatorname{ctg} \frac{i}{2} \frac{dt}{dE} - \omega_3 \end{aligned}$$

$$+ \omega_1 \sin \theta \cdot \operatorname{ctg} \frac{i}{2} - \omega_2 \cos \theta \cdot \operatorname{ctg} \frac{i}{2}$$

Taking

$$\theta - \sigma = +2K - \sigma_0 - \theta_0 - (2\eta + m)$$

$$\theta + \sigma = -2N + \sigma_0 + \theta_0 - (2\alpha + m)$$

into consideration, we obtain

$$\frac{d}{dE} (\theta - \sigma) = +2 \frac{dK}{dE} - (2\eta + m)$$

$$\frac{d}{dE} (\theta + \sigma) = -2 \frac{dN}{dE} - (2\alpha + m)$$

It follows from equations 43 to 46

$$\begin{aligned} \frac{dK}{dE} &= \left( \eta + \frac{m - \omega_3}{2} \right) \\ &\quad + \frac{1}{2} hZr \sin(\nu - \sigma) \operatorname{tg} \frac{i}{2} \frac{dt}{dE} \\ &\quad + \frac{1}{2} (\omega_2 \cos \theta - \omega_1 \sin \theta) \operatorname{tg} \frac{i}{2} \\ \frac{dN}{dE} &= - \left( \alpha + \frac{m - \omega_3}{2} \right) \\ &\quad - \frac{1}{2} hZr \sin(\nu - \sigma) \operatorname{ctg} \frac{i}{2} \frac{dt}{dE} \\ &\quad + \frac{1}{2} (\omega_2 \cos \theta - \omega_1 \sin \theta) \operatorname{ctg} \frac{i}{2} \end{aligned}$$

$$\begin{aligned} \frac{di}{dE} &= hZr \cos(\nu - \sigma) \frac{dt}{dE} \\ &\quad - (\omega_1 \cos \theta + \omega_2 \sin \theta) \end{aligned}$$

From the defining equations

$$\lambda_1 = \sin(i/2) \cos N$$

$$\lambda_2 = \sin(i/2) \sin N$$

$$\lambda_3 = \cos(i/2) \sin K$$

$$\lambda_4 = \cos(i/2) \cos K$$

and from (38), (47), and (48) we obtain

$$\frac{d\lambda_1}{dE} = + \left( \alpha + \frac{m - \omega_3}{2} \right) \cdot \lambda_2$$

$$+ \frac{1}{2} h Z r \cos (\nu - \sigma - N) \cos \frac{i}{2} \frac{dt}{dE} \\ - \frac{1}{2} [\omega_1 \cos (\theta + N)$$

$$+ \omega_2 \sin (\theta + N)] \cos \frac{i}{2} \quad (49)$$

$$= - \left( \alpha + \frac{m - \omega_3}{2} \right) \cdot \lambda_1$$

$$- \frac{1}{2} h Z r \sin (\nu - \sigma - N) \cos \frac{i}{2} \frac{dt}{dE} \\ + \frac{1}{2} [-\omega_1 \sin (\theta + N)$$

$$+ \omega_2 \cos (\theta + N)] \cos \frac{i}{2} \quad (50)$$

$$= + \left( \eta + \frac{m - \omega_3}{2} \right) \cdot \lambda_4$$

$$+ \frac{1}{2} h Z r \sin (\nu - \sigma - K) \sin \frac{i}{2} \frac{dt}{dE} \\ + \frac{1}{2} [-\omega_1 \sin (\theta - K)$$

$$+ \omega_2 \cos (\theta - K)] \cdot \sin \frac{i}{2} \quad (51)$$

$$= - \left( \eta + \frac{m - \omega_3}{2} \right) \lambda_3$$

$$- \frac{1}{2} h Z r \cos (\nu - \sigma - K) \sin \frac{i}{2} \frac{dt}{dE} \\ + \frac{1}{2} [+ \omega_1 \cos (\theta - K)$$

$$+ \omega_2 \sin (\theta - K)] \sin \frac{i}{2} \quad (52)$$

where  $r$  is the radius vector of the real satellite

$$= (1 + \nu) \bar{p} = (1 + \nu)(1 - e_0 \cos E) \quad (53)$$

the 'bar' meaning, as before, a replacement operation.

In forming the orthogonal component of the disturbing force the modified disturbing function must be differentiated only with respect to  $\phi$  and  $\lambda$ , and  $\rho$  must be considered a temporary constant. We designate

$$= \sum_{n=1}^{\infty} \sum_{m=0}^n \left( \frac{a_0}{\rho} \right)^{m+n+1} (1 + \nu)^{-n-1}$$

$$\frac{d^{m+1} P_n(\psi)}{d\psi^{m+1}} \text{grad } \psi \cdot \mathbf{R}$$

$$\cdot (\alpha_{nm} X_m + \beta_{nm} Y_m) \\ + \sum_{n=1}^{\infty} \sum_{m=1}^n \left( \frac{a_0}{\rho} \right)^{m+n+1} (1 + \nu)^{-n-1} \frac{d^m P_n(\psi)}{d\psi^m} \\ \cdot (\alpha_{nm} \text{grad } X_m + \beta_{nm} \text{grad } Y_m) \cdot \mathbf{R} \quad (54)$$

and evidently we have

$$Z = \frac{\partial \Omega}{\partial Z}$$

The rectangular coordinates of the auxiliary satellite are

$$\xi = \rho \cos \phi \cos \lambda \\ \eta = \rho \cos \phi \sin \lambda \quad (55) \\ \zeta = \rho \sin \phi$$

and it follows from

$$\rho^m \cos^m \phi \exp(jm\lambda) = (\xi + j\eta)^m$$

or from

$$X_m + jY_m = a_0^{-m} (\xi + j\eta)^m \quad (56)$$

that

$$\mathbf{R} \cdot \text{grad } (X_m + jY_m) \\ = m a_0^{-1} \sin i (\sin \theta - j \cos \theta) \\ \cdot (X_{m-1} + jY_{m-1}) \quad (57)$$

and

$$\rho \mathbf{R} \cdot \text{grad } X_m = m \frac{\rho}{a_0} \\ \cdot \sin i (X_{m-1} \sin \theta + Y_{m-1} \cos \theta) \quad (58)$$

$$\rho \mathbf{R} \cdot \text{grad } Y_m = m \frac{\rho}{a_0} \\ \cdot \sin i (-X_{m-1} \cos \theta + Y_{m-1} \sin \theta) \quad (59)$$

$$\rho \mathbf{R} \cdot \text{grad } \psi = \cos i \quad (60)$$

Taking equations 37' and 38' into consideration, we deduce

$$\cos i = \lambda_4^2 + \lambda_3^2 - \lambda_1^2 - \lambda_2^2 \quad (60')$$

$$\sin i \cos \theta = 2(\lambda_1 \lambda_4 + \lambda_2 \lambda_3) \cos (\theta) \\ + 2(\lambda_2 \lambda_4 - \lambda_1 \lambda_3) \sin (\theta) \quad (60'')$$



$$\sin i \sin \theta = 2(\lambda_1 \lambda_4 + \lambda_2 \lambda_3) \sin(\theta) - 2(\lambda_2 \lambda_4 - \lambda_1 \lambda_3) \cos(\theta) \quad (60''')$$

Taking equations 58 to 60''' and 54 into consideration we deduce

$$rZ = (1 + \nu) \bar{p} \frac{\partial \bar{\Omega}}{\partial Z} = \sum_{n=1}^{\infty} \sum_{m=0}^{n-1} \left( \frac{a_0}{\rho} \right)^{m+n+1} (1 + \nu)^{-n} \frac{d^{m+1} P_n(\psi)}{d\psi^{m+1}} \cdot (F_{nm} X_m + G_{nm} Y_m) \quad (61)$$

where

$$F_{nm} = \alpha_{nm} \cos i + [\alpha_{n,m+1} \sin i \sin \theta - \beta_{n,m+1} \sin i \cos \theta](m+1) \quad (61')$$

$$G_{nm} = \beta_{nm} \cos i + [\alpha_{n,m+1} \sin i \cos \theta + \beta_{n,m+1} \sin i \sin \theta](m+1) \quad (61'')$$

The coefficients  $F_{nm}$  and  $G_{nm}$  can be represented in the form of trigonometric series using equations 60' and 60''.

In the previous article we deduced the formulas

$$\frac{dt}{dE} = \frac{\bar{p}}{a_0 n_0} \Lambda \quad (62)$$

where

$$\Lambda = \frac{1 - v^2}{1 + \bar{W}} \left( 1 + \frac{y}{\sqrt{1 - e_0^2}} \frac{\bar{p}}{a_0} \right) \quad (63)$$

and

$$n_0^2 a_0^3 = 1 \quad (\text{in Vanguard units}) \quad (64)$$

We also have by definition

$$h_0 = 1/\sqrt{a_0(1 - e_0^2)} \quad (65)$$

Substituting equations 61-65 into 49-52 and taking equations 9 and 6 into account we deduce, after some easy transformations, the following general formulas:

$$\begin{aligned} \frac{d\lambda_1}{dE} = & + \left( \alpha + \frac{m - \omega_3}{2} \right) \lambda_2 \\ & + \frac{1}{2} \frac{h}{h_0} \cdot \frac{a_0 r Z \Lambda}{\sqrt{1 - e_0^2}} (+\lambda_4 l - \lambda_3 m) \\ & + \frac{1}{2} \lambda_3 [\omega_1 \sin(\theta) - \omega_2 \cos(\theta)] - \frac{1}{2} \lambda_4 \end{aligned}$$

$$\cdot [\omega_1 \cos(\theta) + \omega_2 \sin(\theta)]$$

$$\begin{aligned} \frac{d\lambda_2}{dE} = & - \left( \alpha + \frac{m - \omega_3}{2} \right) \lambda_1 \\ & + \frac{1}{2} \frac{h}{h_0} \cdot \frac{a_0 r Z \Lambda}{\sqrt{1 - e_0^2}} (-\lambda_3 l - \lambda_4 m) \\ & - \frac{1}{2} \lambda_{33} [\omega_1 \cos(\theta) + \omega_2 \sin(\theta)] - \frac{1}{2} \lambda_4 \cdot [\omega_1 \sin(\theta) - \omega_2 \cos(\theta)] \quad (66) \\ \frac{d\lambda_3}{dE} = & + \left( \eta + \frac{m - \omega_3}{2} \right) \lambda_4 \end{aligned}$$

$$\begin{aligned} & + \frac{1}{2} \frac{h}{h_0} \cdot \frac{a_0 r Z \Lambda}{\sqrt{1 - e_0^2}} (+\lambda_2 l + \lambda_1 m) \\ & - \frac{1}{2} \lambda_1 [\omega_1 \sin(\theta) - \omega_2 \cos(\theta)] + \frac{1}{2} \lambda_2 \cdot [\omega_1 \cos(\theta) + \omega_2 \sin(\theta)] \\ \frac{d\lambda_4}{dE} = & - \left( \eta + \frac{m - \omega_3}{2} \right) \lambda_3 \end{aligned}$$

$$\begin{aligned} & + \frac{1}{2} \frac{h}{h_0} \cdot \frac{a_0 r Z \Lambda}{\sqrt{1 - e_0^2}} (-\lambda_1 l + \lambda_2 m) \\ & + \frac{1}{2} \lambda_1 [\omega_1 \cos(\theta) + \omega_2 \sin(\theta)] + \frac{1}{2} \lambda_2 \cdot [\omega_1 \sin(\theta) - \omega_2 \cos(\theta)] \end{aligned}$$

where

$$\begin{aligned} l = & \cos^2 \frac{\varphi_0}{2} \cos(E + (\omega)) + \sin^2 \frac{\varphi_0}{2} \cdot \cos(E - (\omega)) - \sin \varphi_0 \cos(\omega) \quad (67) \\ m = & \cos^2 \frac{\varphi_0}{2} \sin(E + (\omega)) - \sin^2 \frac{\varphi_0}{2} \cdot \sin(E - (\omega)) - \sin \varphi_0 \sin(\omega) \quad (68) \end{aligned}$$

Equations 66, being quite general, permit one to take into account also the influence of precession and nutation on the motion of the orbital plane of the satellite.

Until now we have left  $m$  undefined. In order to simplify equations 66 we put

$$m = \omega_3$$

This makes  $\alpha$  and  $\eta$  of the order of the perturbations. They are determined from the condition that the constant terms in the second and third equations of 66 must be absent.

The following classic formulas give the connection between  $\omega_1$ ,  $\omega_2$ ,  $\omega_3$  and  $\vartheta$ ,  $z$ ,  $\zeta$

$$\begin{aligned} &= +\frac{d\zeta}{dE} \sin \vartheta \cos z + \sin z \frac{d\vartheta}{dE} \\ &= -\frac{d\zeta}{dE} \sin z \sin \vartheta + \cos z \frac{d\vartheta}{dE} \quad (67'') \\ &= -\cos \vartheta \frac{dz}{dE} + \frac{dz}{dE} \end{aligned}$$

which include the daily rotation, the precession, and the nutation. Only long-period terms appear in the terms proportional to  $\omega_1$ ,  $\omega_2$ . They are produced by the precession and the nutation. As we can easily see, the precession will produce very small long-period terms with an argument small to the mean node. The choice of long-period terms, to be kept in the development, must be flexible, and it must be compatible with the accuracy to which the potential of the central body is known.

Time is eliminated from the development of  $\omega_1$ ,  $\omega_2$  into trigonometric series by means of Euler's equation. We may think that even the substitution

$$n_0(t - t_0) = \Delta E$$

will suffice to keep under control all significant long-period terms produced by the rotation of the earth.

*Development of the coordinates.* The author [1959] suggested that the coordinates be determined by means of the formula

$$\mathbf{r} = (1 + \nu)\Gamma \begin{bmatrix} a_0 & (\cos E - e_0) \\ a_0 & 1 - e_0^2 \sin E \\ 0 \end{bmatrix}$$

where  $\Gamma$  is a rotation matrix, which can be represented as a product of three matrices. The middle matrix is a matrix rational in  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$ ,  $\lambda_4$ . The numerical values of  $\lambda_i$  for a given time  $t$  were used to obtain the coordinates. The author's present standpoint is that this way has certain advantages if only the numerical values of the coordinates are required. The disadvantage is that we do not see the analytical form of the coordinates and the relative importance of terms in the coordinates. However, it would not take much effort to produce the development of the coordinates in addition to the development of the perturbations. Taking one step more, and developing, we deduce the following formulas that can serve, after the substitution of series for  $1 + \nu$ ,  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$ ,  $\lambda_4$ , to produce the series for the coordinates of the satellite with respect to  $(X'Y'Z')$ .

$$c/a_0(1 + \nu)$$

| $\gamma = iE + j(\omega) + k(\theta)$ |     |     | Coefficients of<br>$\cos \gamma$                          | Coefficients of<br>$\sin \gamma$                  |
|---------------------------------------|-----|-----|---|---|
| $i$                                   | $j$ | $k$ |   |   |
| +1                                    | +1  | +1  | $+(\lambda_4^2 - \lambda_3^2) \cos^2 \frac{\varphi_0}{2}$ | $-2\lambda_3\lambda_4 \cos^2 \frac{\varphi_0}{2}$ |
| +1                                    | -1  | -1  | $+(\lambda_4^2 - \lambda_2^2) \sin^2 \frac{\varphi_0}{2}$ | $+2\lambda_3\lambda_4 \sin^2 \frac{\varphi_0}{2}$ |
| 0                                     | +1  | +1  | $-(\lambda_4^2 - \lambda_3^2) \sin \varphi_0$             | $+2\lambda_3\lambda_4 \sin \varphi_0$             |
| +1                                    | +1  | -1  | $+(\lambda_1^2 - \lambda_2^2) \cos^2 \frac{\varphi_0}{2}$ | $+2\lambda_1\lambda_2 \cos^2 \frac{\varphi_0}{2}$ |
| +1                                    | -1  | +1  | $+(\lambda_1^2 - \lambda_2^2) \sin^2 \frac{\varphi_0}{2}$ | $-2\lambda_1\lambda_2 \sin^2 \frac{\varphi_0}{2}$ |
| 0                                     | +1  | -1  | $-(\lambda_1^2 - \lambda_2^2) \sin \varphi_0$             | $-2\lambda_1\lambda_2 \sin \varphi_0$             |

| $y/a_0(1 + v)$                        |     |     | Coefficients of<br>$\cos \gamma$                  | Coefficients of<br>$\sin \gamma$                          |
|---------------------------------------|-----|-----|---|---|
| $\gamma = iE + j(\omega) + k(\theta)$ |     |     |   |   |
| $i$                                   | $j$ | $k$ |   |   |
| +1                                    | +1  | +1  | $+2\lambda_3\lambda_4 \cos^2 \frac{\varphi_0}{2}$ | $+(\lambda_4^2 - \lambda_3^2) \cos^2 \frac{\varphi_0}{2}$ |
| +1                                    | -1  | -1  | $+2\lambda_3\lambda_4 \sin^2 \frac{\varphi_0}{2}$ | $-(\lambda_4^2 - \lambda_3^2) \sin^2 \frac{\varphi_0}{2}$ |
| 0                                     | +1  | +1  | $-2\lambda_3\lambda_4 \sin \varphi_0$             | $-(\lambda_4^2 - \lambda_3^2) \sin \varphi_0$             |
| +1                                    | +1  | -1  | $-2\lambda_1\lambda_2 \cos^2 \frac{\varphi_0}{2}$ | $-(\lambda_1^2 - \lambda_2^2) \cos^2 \frac{\varphi_0}{2}$ |
| +1                                    | -1  | +1  | $-2\lambda_1\lambda_2 \sin^2 \frac{\varphi_0}{2}$ | $+(\lambda_1^2 - \lambda_2^2) \sin^2 \frac{\varphi_0}{2}$ |
| 0                                     | -1  | +1  | $+2\lambda_1\lambda_2 \sin \varphi_0$             | $+(\lambda_1^2 - \lambda_2^2) \sin \varphi_0$             |

| $z/a_0(1 + v)$                        |     |     | Coefficients of<br>$\cos \gamma$   | Coefficients of<br>$\sin \gamma$   |
|---------------------------------------|-----|-----|--|--|
| $\gamma = iE + j(\omega) + k(\theta)$ |     |     |  |  |
| $i$                                   | $j$ | $k$ |  |  |
| +1                                    | +1  | 0   | $+2(\lambda_2\lambda_4 + \lambda_1\lambda_3) \cos^2 \frac{\varphi_0}{2}$ | $+2(\lambda_1\lambda_4 - \lambda_2\lambda_3) \cos^2 \frac{\varphi_0}{2}$ |
| +1                                    | -1  | 0   | $+2(\lambda_2\lambda_4 + \lambda_1\lambda_3) \sin^2 \frac{\varphi_0}{2}$ | $-2(\lambda_1\lambda_4 - \lambda_2\lambda_3) \sin^2 \frac{\varphi_0}{2}$ |
| 0                                     | +1  | 0   | $-2(\lambda_2\lambda_4 + \lambda_1\lambda_3) \sin \varphi_0$             | $-2(\lambda_1\lambda_4 - \lambda_2\lambda_3) \sin \varphi_0$             |

*Conclusion.* The present method is devised for the use of electronic computers. The basic operation is the multiplication of trigonometric series with numerical coefficients. The multiplication can be performed on the machine automatically, and the terms are retained or rejected on the basis of their numerical values and not on the basis of their order with respect to eccentricity or inclination. For this reason, only the lowest orders in the harmonic analysis need be developed analytically. The higher terms are produced from the lower ones by the application of recurrent formulas and the Horner scheme.

The machine program based on the results of this paper employs an iteration of the perturbations to obtain the actual solution. The solution gives the rectangular coordinates of the position vector in the form of trigonometric series whose arguments are linear in the eccentric anomaly. The method extends to any order

in zonal and tesseral harmonics and therefore will facilitate the search for resonant terms which may only become noticeable after a long period of observation.

The present development of tesseral harmonics may be incorporated in the existing Vanguard program without substantial modifications.

The author wishes to express his gratitude to his colleague Mr. Philip Meyers for his valuable assistance in checking the development.

REFERENCES

Hansen, P. A., *Fundamenta nova investigationis orbitae verae quam Luna perlustrat*, pp. 1-14. Gotha, 1838.  
Musen, P., Application of Hansen's theory to the motion of an artificial satellite in the gravitational field of the earth. *J. Geophys. Res.* 64, 2271-2273, 1959.

# Weather Radar Observations from an Earth Satellite

J. E. KEIGLER AND L. KRAWITZ

*Astro-Electronic Products Division  
Radio Corporation of America  
Princeton, New Jersey*

**Abstract.** The major physical problems encountered in the design of a weather radar satellite stem from the volume of the atmosphere to be sampled and the radar characteristics of clouds, precipitation, and the surface of the earth. Solutions to these problems are sought in the application and extension of conventional radar techniques, electronic scanning, correlation techniques, and precipitation-ground discrimination methods. Consideration is given to the special environment of a satellite radar and to the ground processing of the data to be obtained. The characteristics of a minimal orbital radar system that could be produced today are discussed.

## INTRODUCTION

Contemporary meteorological satellites, such as Explorer VII and the RCA Tiros vehicle, obtain observations by means of radiation detection and television cameras. Lack of knowledge of the actual occurrence and distribution of precipitation is a major deficiency of these observations. At present, radar is the only means by which precipitation detection can be achieved. A satellite-borne weather radar system for overcoming this deficiency must be designed with three major problems in mind in addition to the obvious ones of minimum size, weight, average power consumption: desired coverage area and height; reflectivity and attenuation of both clouds and precipitation at microwave frequencies; and radar ground return from water and land areas.

## DISCUSSION OF THE MAJOR PHYSICAL PROBLEMS

**Coverage.** The area on the earth and the depth of atmosphere above the earth observed by a radar-equipped satellite depend on vertical discrimination desired, the echo detection limitations, and the information-storage and readout capabilities. In order to observe returns on either side of the satellite ground track, a radar beam from the satellite must, in effect, be scanned back and forth in a plane perpendicular to the velocity vector of the satellite. Using straightforward radar techniques of high directivity and echo delay to separate the

returns from a precipitation layer and the ground, the vertical discrimination, or minimum altitude layer that can be resolved, is a function of the satellite altitude, pulse length, antenna beam width, and angle from the nadir, as shown in Figure 1a. At the limiting vertical discrimination,  $\Delta C$ , the leading edge of the radar pulse on the lower side of the beam arrives at the ground at the same time the trailing edge of that pulse on the upper side of the beam passes the radar top of the precipitation layer. This relationship is plotted against the nadir angle for several combinations of the other three parameters in Figure 1b. Immediately evident from these curves is the fact that vertical discrimination is limited much more severely by beam width than by pulse length as the nadir angle increases. Estimates of the meteorological utility of the precipitation data impose an upper limit of approximately 1 mile on the vertical discrimination; deterioration past this limit would leave very little useful information, as most precipitation occurs within the lowest 15,000 feet of the atmosphere. Even with a vertical discrimination of 1 mile some stratiform precipitation would not be observed.

A second factor governing the coverage for which the radar is to be designed is the echo delay, or pulse round-trip time. Since some type of scanning perpendicular to the satellite velocity vector is required, as mentioned above, the time that the antenna points in a given di-



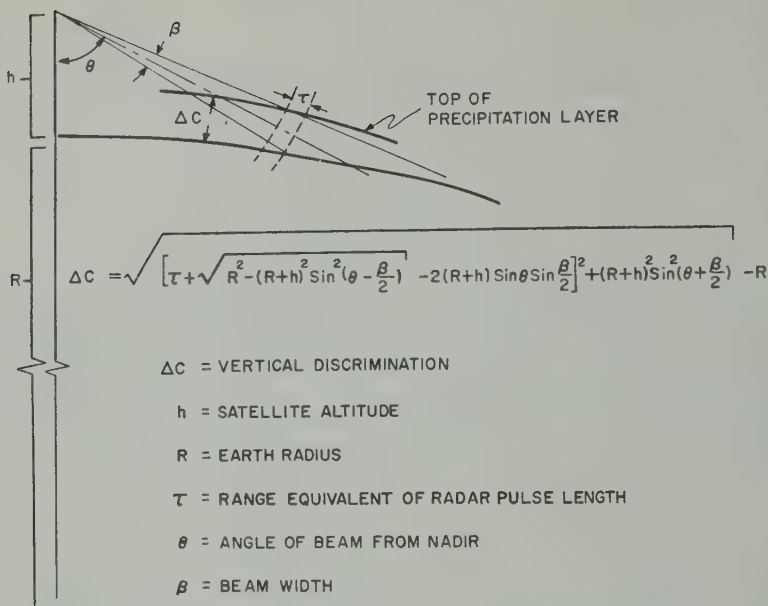
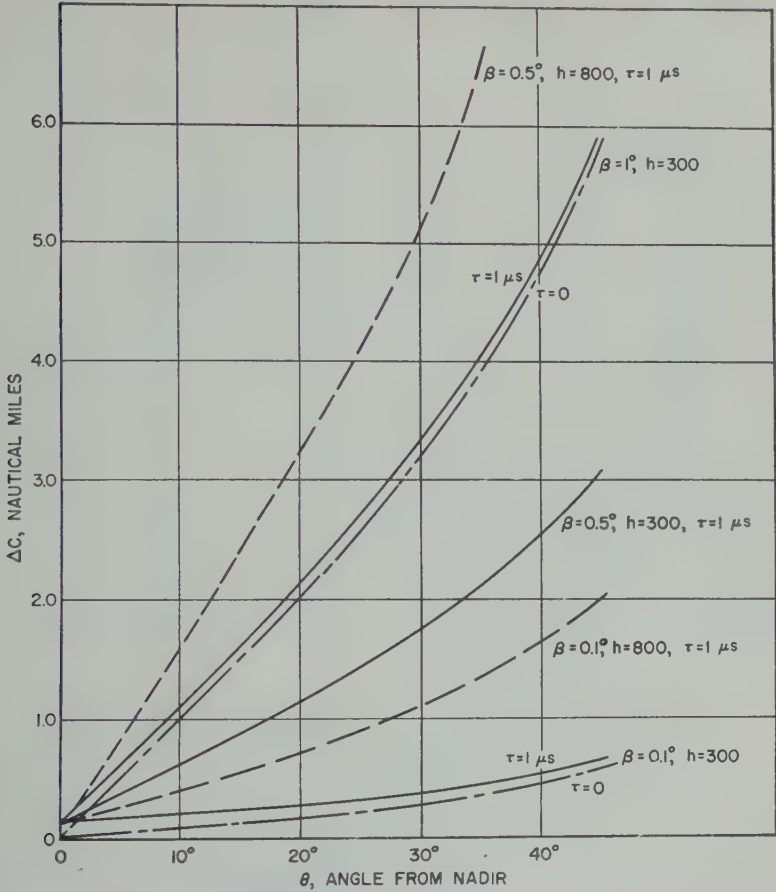


Fig. 1a. Radar beam geometry.

rection is a function of the beam width, the total angle to be scanned, and the satellite altitude; smaller beam widths require more sweeps per unit time to obtain adjacent sweep areas on the earth, and larger total scan angles require greater angular rates per sweep. Assuming a common transmitting and receiving antenna, the time to rotate the antenna through an angle of one beam width (i.e., dwell time per angular resolution element) must be greater than the echo delay for the corresponding resolution element. Figure 2 shows a family of curves, for various satellite altitudes, of the ratio of dwell time per element to echo delay of the maximum slant range at the extremes of the scan. The nadir ground resolution versus altitude and beam-width curves of Figure 3, and the scan angle for complete coverage (scanned strips on successive orbits just meet at the equator) versus altitude curve of Figure 4, are included here as references to give proper perspective to the values of the curves of Figure 2. Thus, for example, Figure 2 indicates that, for a satellite at 300 nautical miles, echo delay time considerations limit the total scan angle to  $35^\circ$ . Figure 4 shows that a total scan angle of  $125^\circ$  is actually required for complete coverage at the equator.

A satellite at an altitude of 1000 miles is precluded from complete coverage at the equator by the echo delay time.

Finally, the information-storage capacity on board the satellite, and the channel capacity of the transmission link from satellite-to-ground stations for readout of this information, also affect the designed coverage area and altitude. The total number of elements of volume sampled per orbit is proportional to the total scan angle and maximum altitude of interest and inversely proportional to satellite altitude, radar pulse length, and the square of the beam width. If the total stored information is read out once per orbit, it can be conveniently expressed as a time-bandwidth product required for complete readout. Assuming a perfect bandwidth-to-time rate ratio of 0.5, the curves of Figure 5 show the readout time-bandwidth product in megacycles-minutes for complete terrestrial coverage at various altitudes. Figure 5 is normalized to a maximum altitude of interest of 1 mile and a radar pulse length of 1 microsecond. The resolution curves of Figure 3 have been superimposed to facilitate translation from ground resolution to information capacity. Figure 5 shows as an example, that for a radar satellite at



1b. Vertical discrimination as a function of vertical angle, beam width, altitude, and pulse length.

ical miles, utilizing a beam width of  $0.5^\circ$  and pulse length of  $1 \mu sec$ , assuming a maximum angle of interest of 1 mile, the nadir ground resolution is about 5 nautical miles and the information to be read out per orbit for complete coverage at the equator is  $(0.18 \text{ Mc-min}) \times 60 \times 2 = 21.6 \times 10^6$  bits.

**Reflectivity and attenuation.** As in any radar system, the design parameters of the meteorological satellite radar will be determined by the size of the target to be detected and the attenuation over the path between the radar and the target. Since the radar cross section of precipitation is a function of its volume rather than its physical cross section or surface area, radar reflectivity is measured in terms of the radar cross section of a perfectly diffuse reflector per unit volume of precipitation, for example  $f^{-1}$  or  $m^{-1}$ .

Thus, if the radar is capable of detecting a target of radar cross section  $\sigma$  at a given distance, and the precipitation fills the radar beam of area  $A$  at that distance, the required thickness of the precipitation of backscatter cross section per unit volume,  $\eta$ , is

$$\Delta d = \sigma / A \eta$$

The backscatter cross section is, of course, dependent on radar frequency and the precipitation rate expressed as depth of precipitation accumulated at the surface of the earth per unit time. Measurements taken at two common frequencies as summarized by Gunn and East [1954] are plotted in Figure 6 for a wide range of precipitation rates, indicating that, from the reflectivity standpoint alone, higher frequencies are preferable.

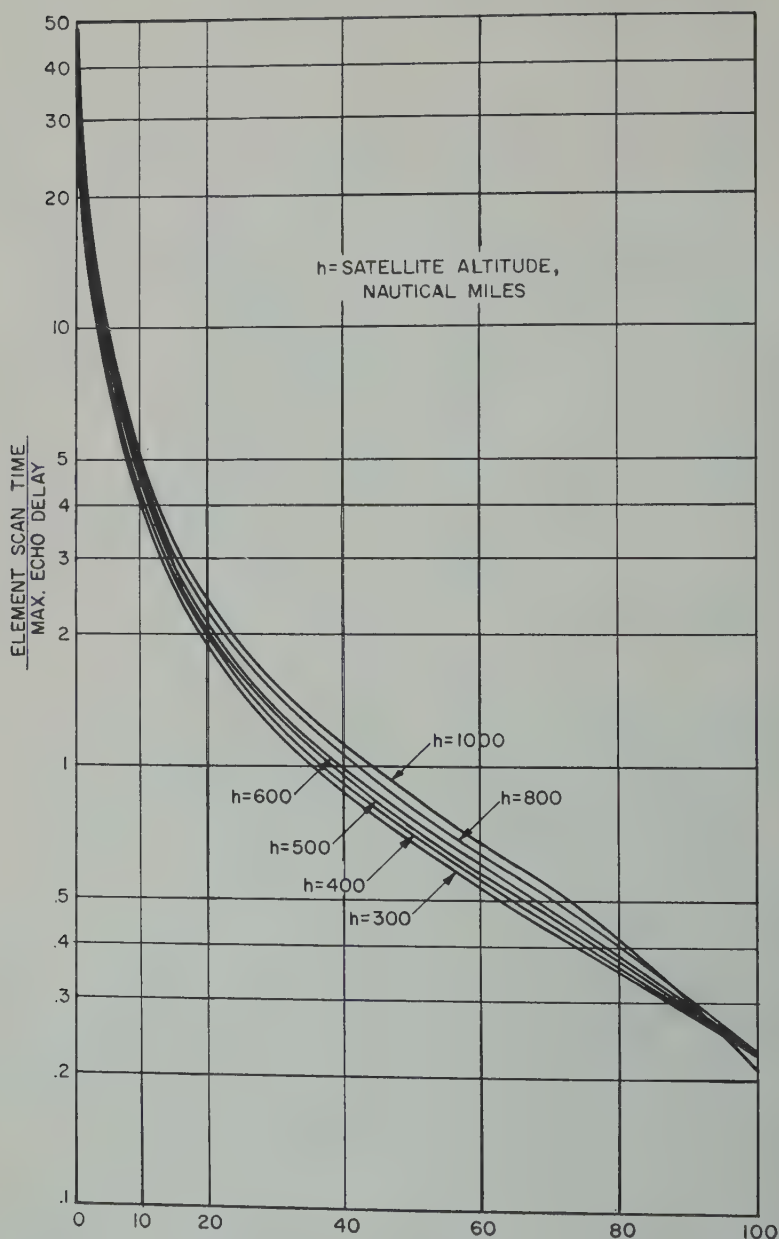
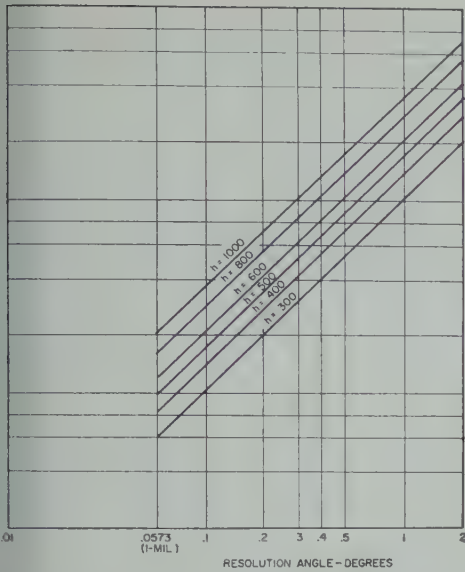


Fig. 2. Elemental scan time to echo delay time ratio for adjacent sweeps of  $1^\circ$  beam (multiply by  $\phi$  for  $\phi^\circ$  beam).

Opposed to this reflectivity advantage of higher radar frequencies is the increased attenuation along the path length with increasing frequency. In addition to the normal frequency dependence of the radar range equation for free space, the two-way attenuation through water

vapor, oxygen, ice clouds, and water clouds in the atmosphere which are between the satellite and the precipitation to be observed is also frequency-dependent, as shown in Figures 7a, 7b, and 7d. Although the water-vapor curve has an absorption line at 1.35 cm, the over-



3. Nadir ground resolution as a function of altitude and resolution angle.

attenuation through clouds, oxygen, and water vapor in any reasonable depth of atmosphere remain a monotonically increasing function of frequency both because of the increasing attenuation by water clouds at higher frequencies because of the decreasing water-vapor content of the atmosphere with altitude. Thus, the choice of operating frequency for optimum performance need not be too strongly influenced by the water-vapor line at 1.35 cm.

It should also be mentioned that the attenuation by ice clouds, water clouds, and rainfall is temperature-dependent. The attenuation increases with increasing temperature in ice clouds and decreases with increasing temperature in water clouds. In ice clouds the effect is a factor of about 5 in 20°C, and in water clouds, a factor of about 2 in 20°C.

**Ground return.** Unlike conventional ground-based or airborne weather radars which look parallel to the ground away from the earth, the satellite radar looks nearly vertically down on the precipitation on the ground immediately below, the return being from the precipitation actually extending into the ground return. Since the vertical extent of precipitation is desired, this ground return provides a useful reference for measurement of the relatively short distance of the precipita-

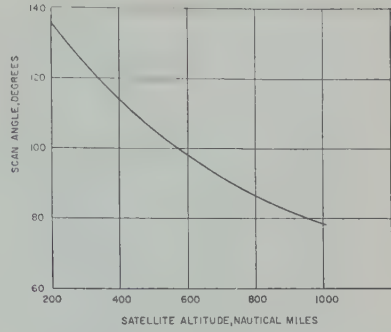


Fig. 4. Scan angle for complete coverage at equator.

tion above the earth to avoid measuring the much longer distances from the satellite. Whereas, for precipitation detection, it is precipitation reflectivity and attenuation through the region above it that are pertinent, as discussed in the preceding paragraph, detection of ground return for use as a reference involves the ground backscatter cross section and attenuation through the region above it, namely clouds plus precipitation. Attenuation per mile through precipitation is much greater than through clouds, as shown in Figures 7d and e, again increasing monotonically with frequency. The attenuation by rain within the Rayleigh region, that is, for wavelengths 10 cm or longer, roughly doubles from 18° to 0°C. At wavelengths shorter than 10 cm the temperature effect on attenuation by rain is wavelength-dependent. The attenuation decreases by a factor of about 30 per cent from 18° to 0°C at wavelengths near 3 cm [Ryde, 1946].

Table 1 lists the total vertical radar attenuation at several frequencies through the lower 12 miles of the atmosphere for various meteorological conditions [Atlas and others, 1952]. There is virtually no attenuation above this level.

Again unlike ground or airborne radars, the 'target' on the far side of this attenuating region in this case is an immense one, the earth. It fills the entire beam and has a reflectivity several orders of magnitude greater than the other targets that the radar must be designed to detect, namely precipitation. The ground backscatter cross-section per unit area varies widely over various types of terrain and sea states. Values ranging from 0.002 for deserts to 0.6 for dense metropolitan areas have been



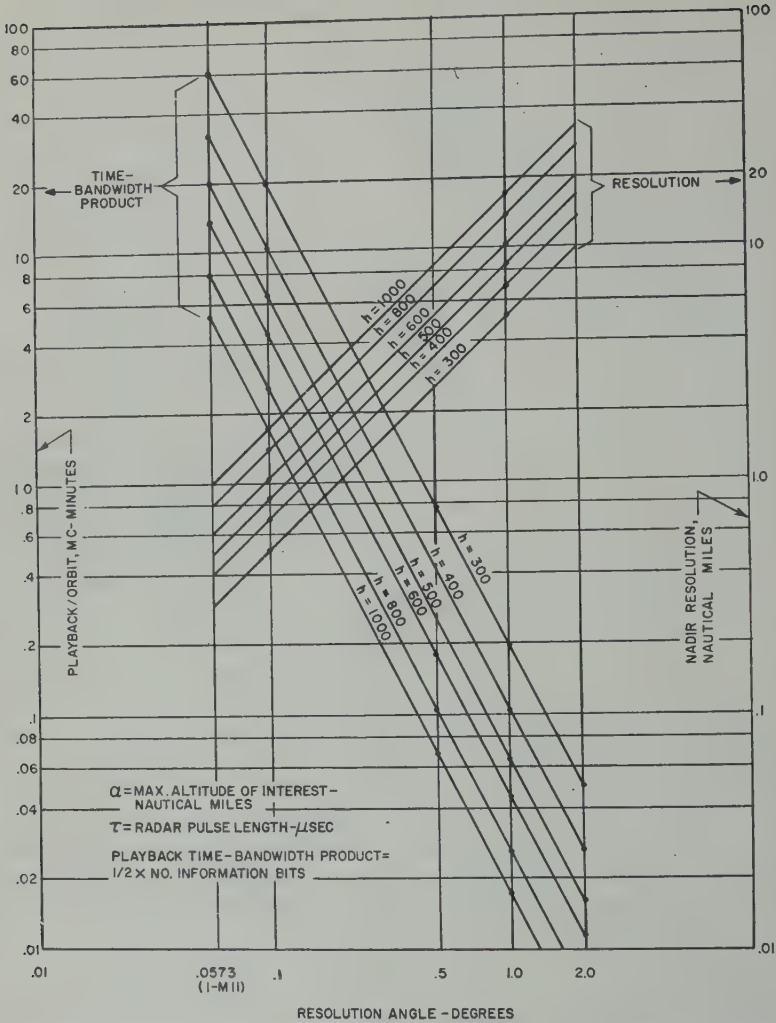


Fig. 5. Information per orbit for total coverage at equator (multiply playback values by the ratio  $\alpha$  in miles to  $\tau$  in microseconds).

measured at the extremes. These values of backscatter cross section for terrain are virtually independent of frequency and angle of incidence of the beam. Measurements over water have indicated a wider range primarily due to the exceptional case of specular reflection. In the rare instance when the water surface is smooth compared with a radar wavelength over the entire area covered by the beam, specular reflection causes the backscatter cross section to exceed unity at vertical incidence and drop nearly to zero for all angles away from the vertical. For the much more probable conditions of rough

(compared with a wavelength) sea surfaces backscatter cross section ranges from 0.01 to 1.0 for various sea states independent of frequency and increasing with angle of incidence.

POSSIBLE APPROACHES TO A SATELLITE RADAR

It is not apparent that any single approach to this problem is completely satisfactory. Conventional radar techniques, electronic scan correlation techniques, and precipitation-gate discrimination methods are all worthy of consideration.

Application and extension of conventional

iques. As indicated by the vertical dis-  
nation curves of Figure 1b, a conventional  
employing only beam directivity and echo  
to obtain resolution will fulfill the dis-  
nation requirement even at large nadir  
s provided that a sufficiently narrow beam  
employed. Beam width and gain relationships  
parabolic antennas as functions of antenna  
and radar frequency are shown in Figure  
combination of high frequency and large  
na diameter is required to realize the nar-  
beam (keeping in mind that the selection of  
ency must also be based on reflectivity and  
uation considerations).

TABLE 1. Total Radar Attenuation in Decibels  
along a Vertical Path in the Lowest 12 Miles of  
Atmosphere for Various Synoptic Conditions

| Case |      |      |      |      |      |    |     |
|------|------|------|------|------|------|----|-----|
| Case | I    | II   | III  | IV   | V    | VI | VII |
| 0.20 | 0.20 | 0.19 | 0.40 | 0.40 | 0.50 | 7  |     |
| 0.35 | 0.40 | 0.28 | 0.80 | 0.80 | 2.30 | 22 |     |
| 1.28 | 1.79 | 0.78 | 2.00 | 2.70 | 5.80 | 42 |     |
| 0.73 | 0.84 | 0.63 | 3.00 | 2.75 | 7.70 | 66 |     |

Explanation of cases

- U. S. standard atmosphere (maritime polar mass).
- Maritime tropical air mass.
- An ice cloud, 3 miles deep, of moderate con-  
tation, ice content = 0.2 g/m<sup>3</sup>; maritime  
air overrides modified continental polar air.
- Same as case III, with heavy snow (5 mm/hr)  
rest 2 miles of atmosphere with a 3-mile-thick  
ice-water cloud above. In the cloud, the  
water content is 0.1 g/m<sup>3</sup>, the ice content  
m<sup>3</sup>.
- A 3-mile-thick water cloud (0.2 g/m<sup>3</sup>) in a  
ime polar air mass.
- A continuous rain of moderate intensity 4.5  
r extends up to 2 miles. Above the rain,  
snow is falling at the rate of 4 mm/hr of  
alent water. The snow concentration is 1.2  
A 2-mile-deep water cloud exists, having a  
water content of 0.2 g/m<sup>3</sup> in its lower half  
f 0.05 g/m<sup>3</sup> in its upper half.
- An intense thunder storm in maritime  
al air. Rain is falling in a 3-mile layer at the  
of 25 mm/hr. A thick water cloud exists, 5  
deep, with a liquid water content of 0.5 g/m<sup>3</sup>.  
e the water cloud, a 2-mile-deep ice-water  
exists with an ice concentration of 0.4 g/m<sup>3</sup>  
liquid water content of 0.1 g/m<sup>3</sup>.

Present high-power magnetrons are largely  
confined to X-band frequencies (10,000 Mc/s)  
and below; but several magnetrons with peak  
powers well above 50 kilowatts have recently  
been developed at 24,500 and 35,000 Mc/s.  
These magnetrons are in general limited to very  
short pulse lengths, so that the product of peak  
power by pulse length which is important from  
the standpoints of detection and signal-to-noise  
ratio tends to decrease with increasing frequency.  
However, the techniques of using a train of  
short pulses and interleaving the transmitted  
and received pulses are promising for the satellite  
radar, since, with its long echo delay, additional  
pulses can be transmitted before the first one  
returns without incurring any ambiguity.

In this unique radar where all targets, both  
precipitation and earth, fill the radar beam, the  
returned pulse will always be stretched longer  
than the transmitted pulse. The video range in-  
formation will thus consist of more low-fre-  
quency components than it would for a dis-  
crete target, indicating the possibility of a nar-  
rower bandwidth receiver than would normally  
be associated with a conventional radar of the  
same pulse length. Since the signal-to-noise ratio  
improves directly with decreasing receiver band-  
width (provided that the useful signal is con-  
tained in the reduced band), the optimum  
receiver bandwidth-pulse length combination  
should be selected for this special case.

Long-life satellites of the near future will  
very likely be more limited by available power  
than by orbital weight. Thus, so-called brute-  
force methods of using physically large antennas  
must not be immediately rejected in compari-  
son with more sophisticated methods of signal  
processing (to obtain equivalent resolution)  
which involve more complex circuitry and re-  
quire more operating power. Studies sponsored  
by the Wright Air Development Center have  
been published on the subject of antennas which  
can be unfurled to considerably more than the  
5-foot-diameter limitation of the present launch-  
ing-rocket nose cones. With the weightless con-  
dition obtaining in orbit where the antenna  
would be erected, and the absence of drag after  
erection, a large antenna is indeed promising.  
Its strength and rigidity requirements are de-  
termined only by the forces and moments due  
to initial launching of the satellite; also, its

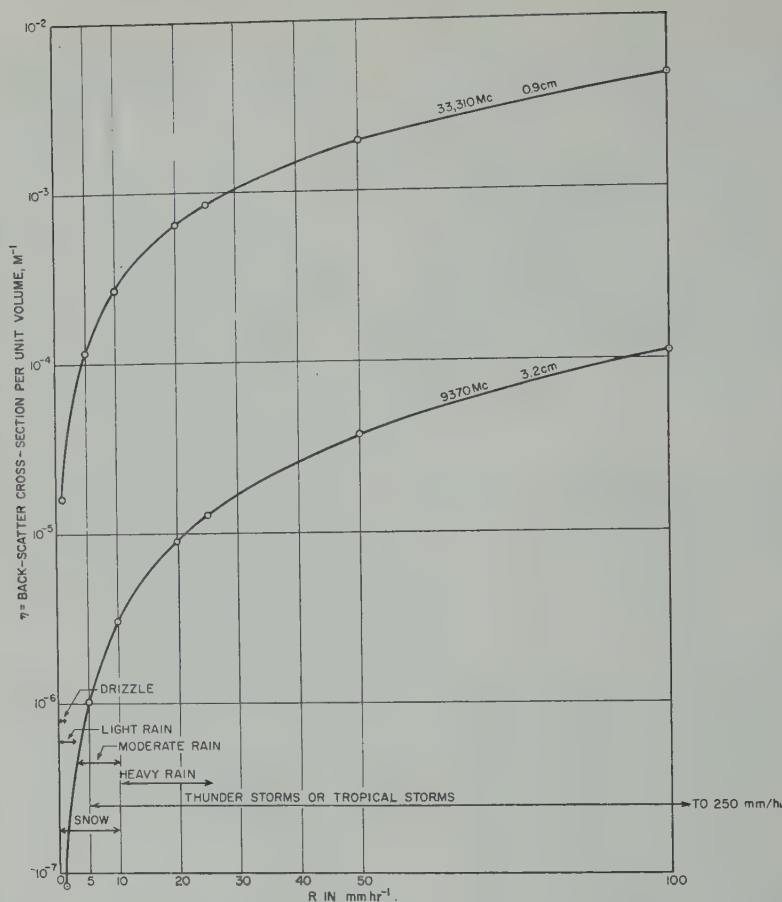


Fig. 6. Backscatter cross section of rain vs. rate of rainfall.

fold-out mechanism power is very low in the absence of gravity.

For narrow-beam antennas, the sweep frequency must be high to cover adjacent sweep areas as the satellite travels at its high orbital velocity. At an altitude of 300 nautical miles and a beam width of  $1^\circ$ , for example, the sweep frequency must be 1 per second. This frequency is inversely proportional to beam width and altitude. To avoid the high power required to oscillate an antenna over an appreciable angle at this rate, a continuously rotating antenna array might be employed. The radar transmitter-receiver would be switched synchronously to each of the antennas, as it scanned the earth. With no air drag present, the only antenna drive power required would be that needed to overcome bearing friction. This arrangement has the advantages

of a reduced rotational frequency, parallel-line scans rather than the zigzag pattern of an oscillating antenna.

Another technique is the possible elimination of scanning by utilizing a fan-shaped beam for transmitting and a many-fingered beam for receiving. A receiver associated with each one of the fingers of the receiving beam will then furnish the necessary directional information.

**Electronic scanning.** The various methods of electronic scanning of the radar beam must be weighed against the physical beam-scanning methods. Hybrid scanning techniques in which there is mechanical motion of some element other than the entire antenna reflector are worthy of consideration. The objective here is to minimize the mass of the moving parts to reduce the perturbing moments that must be over-

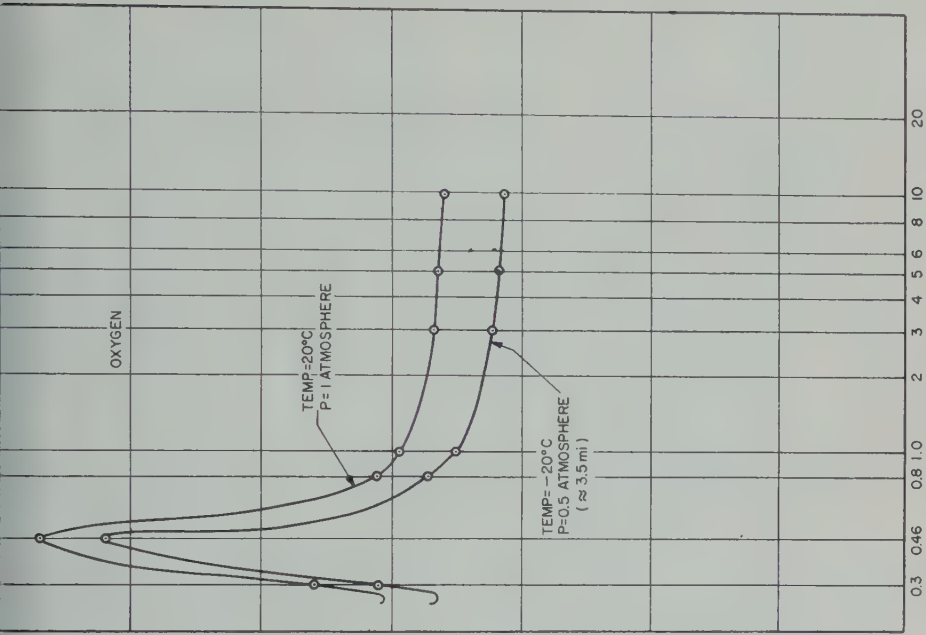


Figure 7b

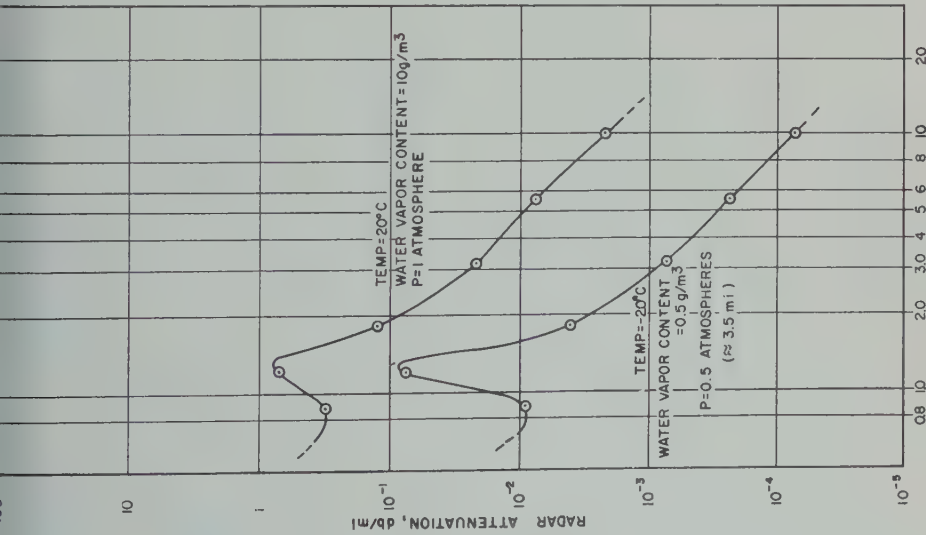


Figure 7a



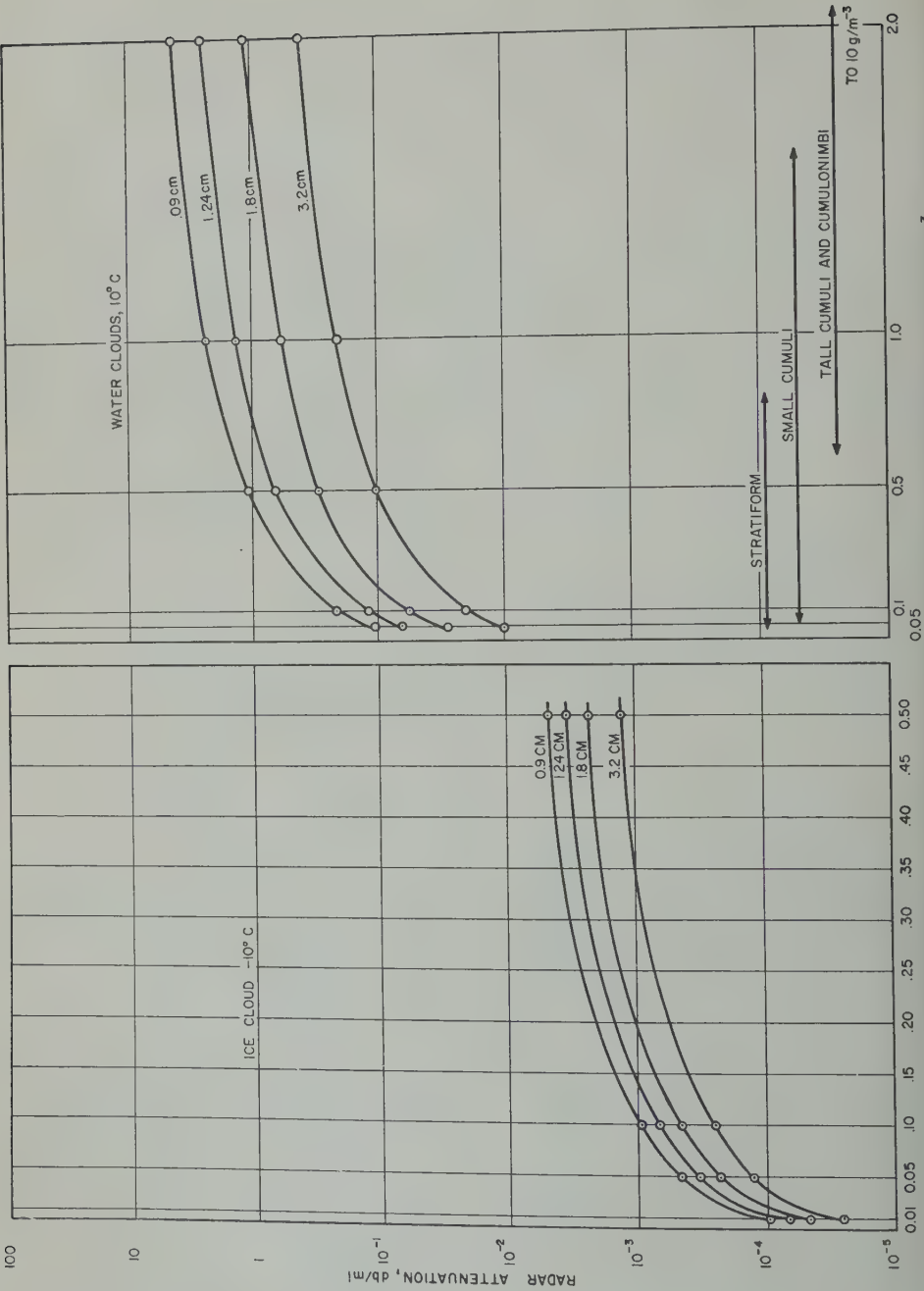


Figure 7d  
E-1335-72

Figure 7c

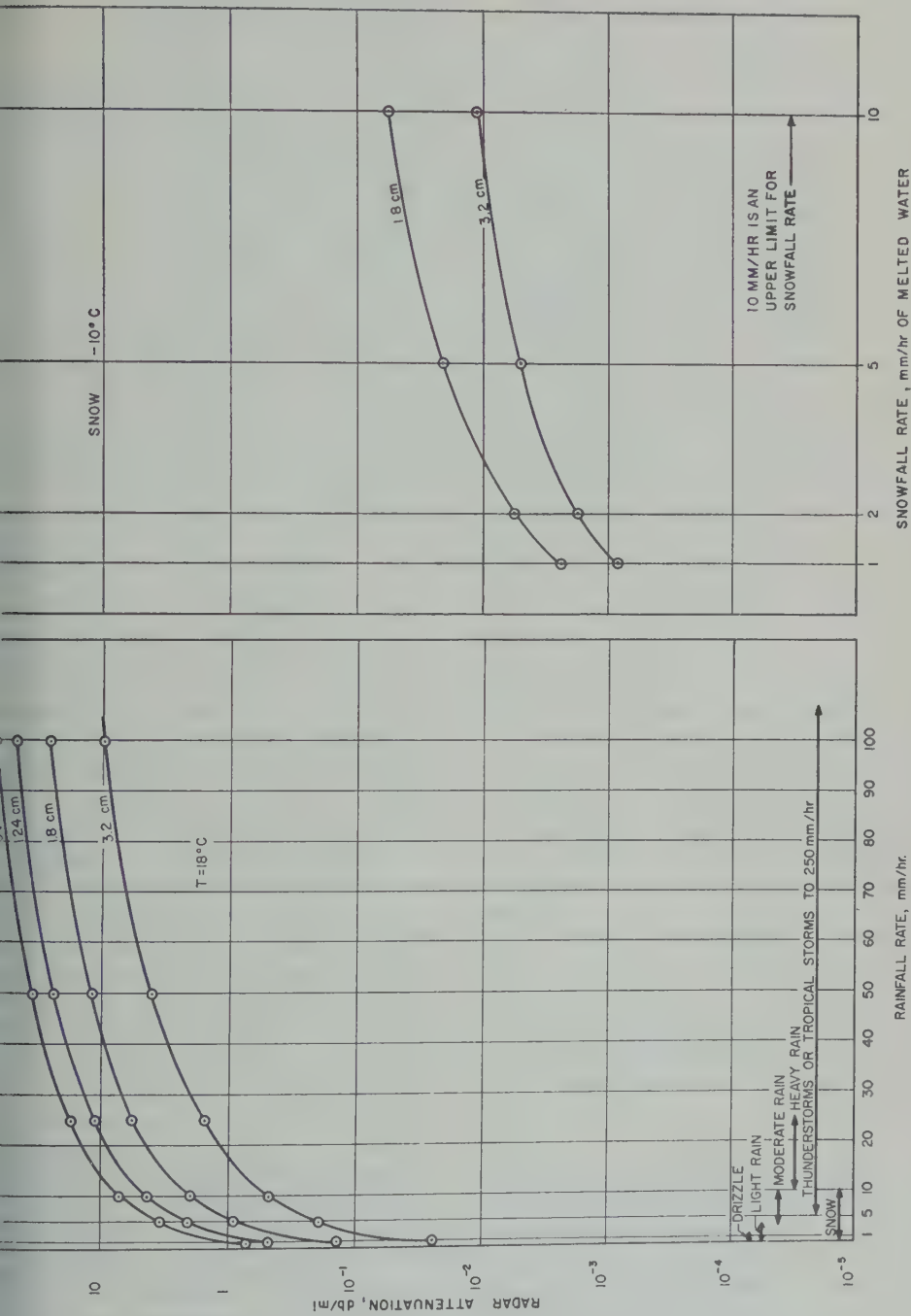


Fig. 7. Frequency dependence of two-way attenuation: (a) through water vapor; (b) through ice cloud,  $-10^\circ\text{C}$ ; (c) through water cloud,  $+10^\circ\text{C}$ ; (d) through rainfall,  $+10^\circ\text{C}$ ; (e) through snowfall,  $-10^\circ\text{C}$ .

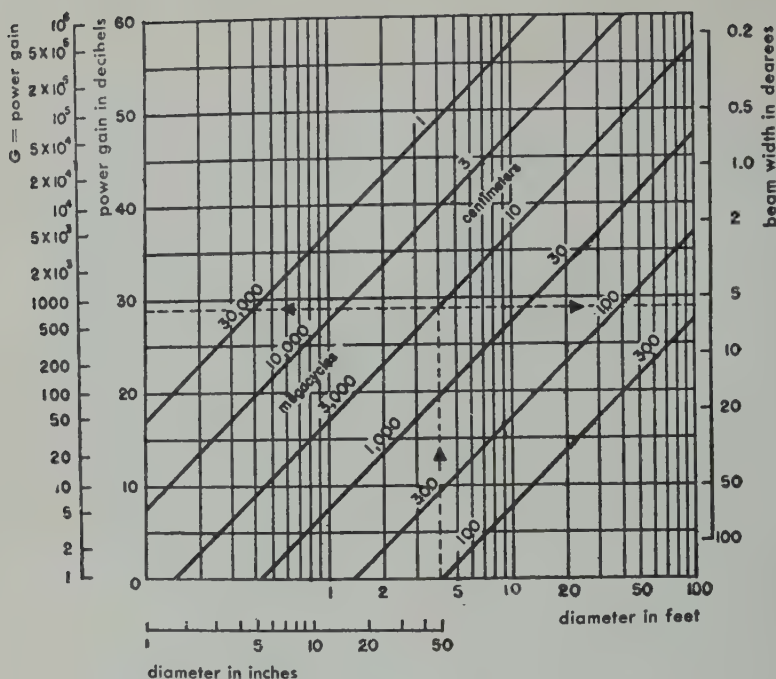


Fig. 8. Beam width and gain of a parabolic reflector.

come by the stabilization system of the satellite.

The purely electronic, or so-called electrical, methods of scanning permit rapid scanning, the rate being limited by the inertia, or time constant, of electrical circuits rather than by mechanical mass. Electrical scanning includes various schemes such as frequency modulation and variation of the dielectric constant, the permeability, and the conductivity of gases. A few attempts at frequency scanning have been reasonably successful at lower frequencies and at the expense of considerable circuit complexity. Recently the phase-shifting characteristics of ferrite materials have appeared to offer some hope of being useful in an antenna array for scanning.

Various hybrid scanning techniques have been developed in the past and have proved successful in rapid scanners for experimental radar systems, both ground-based and airborne types. One airborne system developed by RCA scanned a  $\frac{1}{3}^\circ$  pencil beam through a  $45^\circ$  angle at a rate up to 30 scans per second.

*Correlation techniques.* With vertical discrimination depending so critically on antenna

beam width as shown in Figure 1b, it seems reasonable to seek methods for realizing very high angular resolution without physically large antennas.

One of these methods that has been extensively developed is the monopulse radar with two-lobed antenna pattern. In monopulse air-to-ground ranging systems, accurate range-to-ground is measured along the crossover line of the two lobes by comparing the signals received on them for the condition of equality. The angular accuracy of the lobe crossover sight line is much greater than that of the individual lobes.

Increased angular resolution can also be obtained from the difference in variation in Doppler frequency with time in a direction parallel to the flight path of the radar. This technique developed under Project Michigan makes use of the motion of an aircraft to produce a map of the ground with a side-looking radar. It utilizes range measurement to provide resolution perpendicular to the line of flight. This technique cannot be used directly in the satellite meteorological radar because

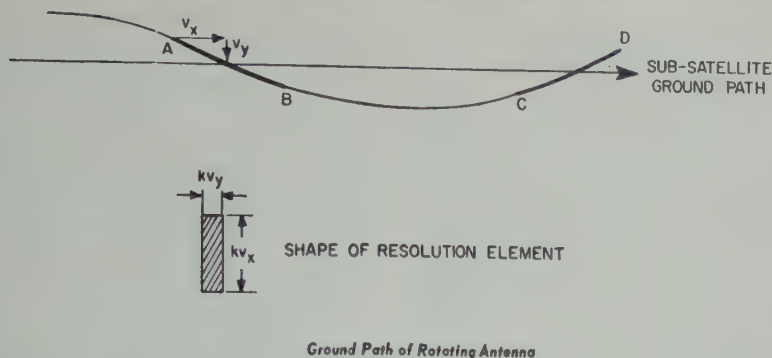


Fig. 9. Ground path of rotating antenna.

resolution at the nadir would be very low, the method being effective only at a large angle from the vertical. It also assumes that the ground is flat and does not provide a three-dimensional picture.

The solid-angle resolution Doppler system was developed to overcome the two-dimensional limitation of the Michigan system. Dr. R. Wilmotte of the Advanced Military Systems group of RCA has designed a receiving antenna that would revolve around a satellite at the end of a cable 40 feet long, with the transmitting antenna at the center of this circle. It is assumed that the simplest and most accurately maintained motion for the receiving antenna is one in which the plane of rotation remains substantially unchanged in space, which means that it rotates relative to the earth. Objects on the ground will therefore have a motion relative to the antenna which will appear as a Doppler frequency in the echo signal reaching the antenna. Since the unique Doppler frequency is immediately known for each point being surveyed, it is then possible to resolve adjacent

The ground path of the antenna projected on the ground will be sinusoidal as shown in Figure 9. The velocity of the satellite will be greater than the circumferential velocity of the antenna, so the resolution in the direction of the satellite will be greater than that in a direction perpendicular to it, in the ratio of the component velocities. It will be seen that the maximum resolution is perpendicular to the line of flight occurring during the portions AB and CD of the path.

If during these portions the velocities are  $V_s$  and  $V_a$ , respectively, the shape of a resolution element will be as shown. The degree of resolution will be substantially equal to that which would be obtained with a single pulse using a receiving antenna having the dimensions of the rectangular area projected onto the ground which encloses the path traveled by the antenna during the time of integration.

*Precipitation-ground discrimination.* In addition to investigations of the feasibility of improving vertical discrimination by improving the angular resolution, possible methods of discriminating between the echoes from ground and precipitation within the same elemental volume (beam width times pulse length) must also be considered. Such methods are necessarily based on one or more of the physical differences between precipitation, and hence of some parameter of its radar echo, and ground (i.e., terrain and water).

The first obvious difference is spatial separation. Even though the precipitation and ground are within the same elemental volume, increased resolution along the pulse-length dimension of the volume is attainable by utilizing leading-edge detection as is done in precision-tracking and fire-control radars rather than simple amplitude detection of the returned pulse. To be weighed against this advantage of improved range resolution is the fact that a wide-band receiver is required for leading-edge detection, thus decreasing the signal-to-noise ratio. The degree of improvement of vertical resolution possible by leading-edge detection is indicated



by the zero-pulse-length curves of Figure 1b; the narrower the beam width, the greater is the percentage improvement.

Precipitation particle size relative to a wavelength causes its backscattering cross section to be a function of radar frequency as shown by the curves of Figure 6, whereas the heterogeneous ground return is independent of frequency. Hence simultaneous observation of a given elemental volume at two frequencies would indicate the presence of precipitation in that volume proportional to the difference in returned signals at the two frequencies.

Particle shape of the precipitation is unique compared with the random ground surface, the particles usually being nearly spherical. These homogeneous spherical surfaces reflect radar energy without any change in the type of polarization but with a change in sense for circular polarization. Elimination of rain clutter on air search radars utilizes this principle. By the same token, a difference in returned signal for linearly (ground plus precipitation) and circularly (ground only) polarized signals could serve as an indication of the presence of precipitation.

Since the precipitation has a vertical velocity relative to the ground, return from it has a Doppler frequency shift. With the continued improvements in narrow-band filtering techniques, the use of continuous-wave or pulse-Doppler radar techniques is worthy of consideration in order to separate in the frequency domain signals that are coincident in the time domain. In addition to its precipitation discrimination possibilities, a Doppler radar could double as the sensor unit for satellite stabilization to the local vertical.

Lastly, because of surface scattering differences of the ground and precipitation integrated over the entire elemental volume, the composition of the video spectrum from the two is different. Observations from the RCA AVQ-10 weather radar show precipitation as a fuzzy-edged area, owing to the poorly defined border of a precipitation area, whereas ground return is sharp. Analysis has corroborated these observations, with the ground video containing more high-frequency components than the precipitation. A related technique for investigation makes use of the Fourier transform of the video time

series in an autocorrelation detector. This technique would be feasible if it could be shown that there is some expectable time homogeneity in the return from either ground or precipitation but not both. Then the stationarity, or lack thereof, of the autocorrelation function would be an indication of whether or not precipitation contributed to the returned signal.

#### SPECIAL ENVIRONMENTAL CONDITIONS

A radar system installed in a satellite must operate under conditions far different from those encountered by any existing radars. Some of those conditions pose major problems that must be overcome; others present unique advantages. As in the design of any satellite equipment, the ability to operate in a vacuum environment is obviously a major problem.

One aspect of the satellite environment that might be utilized to good advantage is the near-perfect vacuum existing at satellite altitudes. Magnetron failure, the cause of perhaps half of all conventional radar malfunctions, is usually the result of cathode poisoning and subsequent deterioration as outgassing from components within the envelope slowly softens the vacuum. By using the infinite-capacity vacuum pump of outer space surrounding the satellite, the magnetron envelope could be opened to the environment once the satellite had reached orbital altitude. There would be continual pumping by this space pump, at a much lower pumping pressure than is attainable in an earth-bound laboratory, which should greatly prolong the magnetron life past the typical 1000 operating hours found in present applications.

Similarly the vacuum environment would obviate the necessity for pressurization of microwave plumbing and other high-voltage sections of the transmitter that is required in airborne radars. This extremely low pressure is far beyond the minimum of the voltage breakdown curve, the breakdown potential, in fact exceeding that at atmospheric pressure.

An accompanying penalty of the vacuum environment is the lack of a convenient cooling medium. Instead of the customary forced-air cooling of the magnetron, intense local heating has to be conducted either by a circulating coolant or by a metallic conductor to a

on the satellite for subsequent dissipation of radiation alone.

### GROUND DATA HANDLING

ly consideration must be given to the ultimate use of satellite radar data by meteorologists. The data-handling process might actually be on board the satellite. Since only the lower layers or so of the atmosphere are of interest, transmission of the video information corresponding only to this small part of the total sweep would be desirable. Otherwise, transmission of the video corresponding to the entire 1000-mile total range would require a larger channel capacity than is indicated by curves of Figure 5. Along with the video, the dependent variables of time and nadir angle must be transmitted to provide information on the location of the satellite and the direction of the beam.

ditional-electronic ground-data-handling equipment can then be specified for converting video, time, and nadir-angle information into successive (in time) vertical sweeps to continuous (in space) presentation, i.e., maps, corresponding to given altitude levels. A position-analog method might be to present the total sweep on a cathode-ray tube as an inclined line, corresponding to the last 10 miles or so of the slant-range path, moving across the tube at a rate and inclination angle given by the nadir angle for that sweep and having one end on a base line corresponding to the earth. A slit camera focused along a slit parallel to the base line would generate a map, corresponding to the level of the slit above the base line on continuously moving film synchronized with the nadir angle sweep rate. A refinement of this technique would be to replace the slit by a series of fiber optics; then several rows of fibers could simultaneously generate maps corresponding to various altitude levels on parallel strips of a single synchronously moving film.

### A MINIMAL ORBITAL RADAR SYSTEM

For the operation of a radar in a satellite to require significant advances in the state of the art, the adaptation of radar design and concepts to orbital conditions is of equal im-

TABLE 2. Characteristics of Minimal Orbital Radar System

|                                |   |
|--------------------------------|---|
| Frequency                      | 10,000 Mc/s                                       |
| Peak power                     | 80 kw   |
| Pulse width                    | 5 $\mu$ sec                                       |
| Duty cycle                     | $10^{-6}$   |
| Receiver noise figure          | 10 db   |
| Detectable radar cross section | $2 \times 10^5$ ft <sup>2</sup><br>(at 800 miles) |
| Average power                  | 36.0 watts  |
| Antenna                        | 4' parabolic,<br>1.75° beam                       |
| Weight, excluding antenna      | 50 lb   |

portance to the optimization of the various parameters as may be determined by future study programs. Such adaptation has been proposed by RCA in the form of a simple precipitation sounding radar. Although this sounding radar would provide only limited meteorological information, it would furnish extremely valuable engineering information for design of subsequent wide coverage precipitation radars. A successful minimal satellite radar would greatly accelerate development of advanced systems utilizing some of the techniques mentioned here.

The nonscanning antenna beam of this radar would require no moving parts or sophisticated circuitry, and the straight pulse amplitude detection receiver would minimize circuit complexity in that part of the system. This simplicity would ensure maximum reliability and make precise evaluation of the radar performance possible. Furthermore, the power requirement would be well within the capability of contemporary solar power supplies. The weight and size including antenna would be within the capacity of contemporary launching vehicles. The feasibility of this sounding radar has already been thoroughly investigated; it could be fabricated and placed in orbit now. The characteristics of such a minimal radar are summarized in Table 2.

A radar satellite yielding full global coverage of precipitation at several levels will not be available in the immediate future. It is hoped that this paper has served to clarify and exemplify the engineering and technical difficulties that must be surmounted to make it possible.

## REFERENCES

- Atlas, D., and others, Weather effects on radar, *Air Force Surveys in Geophysics no. 23*, Air Force Cambridge Research Center, 1952.
- Gunn, K. L. S., and T. W. R. East, The microwave properties of precipitation particles, *Quart. J. Roy. Meteorol. Soc.*, 80, 522, 1954.
- Ryde, J. W., The attenuation and radar produced at centimeter wavelengths by various meteorological phenomena, *Meteorological Factors in Radio Wave Propagation*, The Physical Society, London, 1946.

(Manuscript received April 14, 1960; revised 20, 1960.)

## On the Nonseasonal Variations in Sea Level along the West Coast of North America<sup>1</sup>

GUNNAR I. RODEN

*Scripps Institution of Oceanography  
La Jolla, California*

**Abstract.** Spectra of monthly sea level anomalies are investigated for the frequency range between 0 and 6 cycles per year. It is found that most of the power is concentrated at low frequencies. There is a good coherence between anomalies of sea level and atmospheric pressure at almost all frequencies, indicating a fast response between these variables. The coherence between anomalies of sea level and sea surface temperature is moderate to poor depending upon the particular environment of the tide gage station. There is some coherence between anomalies of sea level and the south component of the geostrophic wind, owing to the influence of the Coriolis force. In river mouths the anomalies in sea level largely reflect the anomalies in river discharge. A fairly successful estimate of the sea level anomaly of the present month can be made from the present anomalies of atmospheric pressure and sea surface temperature. Statistical prediction of sea level anomalies is less successful.

**Introduction.** For many years careful records have been kept of the sea level at various tide gage stations around the periphery of the Pacific and on a few ocean islands. The mean seasonal variation of such records was studied by Nomitsu and Okamoto [1927] and by La Jolla [1939] and Patullo, Munk, Revelle, and Thompson [1955]. Day to day variations in sea level were investigated by Groves [1956, 1957] and Shoji [1957], and a brief study of large irregular variations was made by Stewart, Thompson, and Taylor [1958]. Recently Haubrich and Munk [1959] made a statistical study of sea level in connection with the pole tide. The main conclusions that can be drawn from the above-mentioned investigations are that (1) the mean annual variation of sea level shows a more or less close relation to the mean annual variations in atmospheric pressure and the average heat content of the water, (2) the response of sea level to atmospheric pressure is instantaneous and inverse for periods larger than a day, (3) the variations in sea level are coherent over distances of several hundred kilometers, and (4) there are no periodicities in the spectra.

In this paper the regular seasonal variation is investigated by taking differences between the

monthly mean and long-term mean for the same month. The resulting anomalies are then used to compute the autocorrelations and the cross correlations and from them the power spectral densities, coherence, and phase, using Fourier transforms. In performing these calculations it has been assumed that the anomaly records are stationary, at least in the wide sense, so that the correlation functions depend only upon the correlation interval and are not bound to any origin in time.

The location and environment of the tide gage stations together with some information on the statistics of the various records are given in Table 1 and Figure 1. It is seen that most tide gage stations are located in harbors and bays and that there are only very few stations along the open coast. The tide gage station at Astoria is located in the mouth of the Columbia River; this station has been included in order to show the effect of river discharge upon the sea level.

The length of the sea level records used varies from 72 months to 444 months. The particular length chosen was based upon the availability of simultaneous records of sea level, sea level atmospheric pressure, and sea surface temperature. The number of spectral estimates or correlation intervals was chosen so that the degrees of freedom,  $\nu$ , defined by  $\nu = 2n/m$ , where  $n$

<sup>1</sup>Contribution from the Scripps Institution of Oceanography, New Series.



TABLE 1. Location and Environment of Tide Gage Stations ( $n$  = total record length in months;  $m$  = number of spectral estimates;  $\nu$  = degrees of freedom)

| Tide Gage Station  | N-Lat. | W-Long. | Environment   | $n$ | $m$ | $\nu$ |
|--------------------|--------|---------|---------------|-----|-----|-------|
| Hawaii             |        |         |               |     |     |       |
| Honolulu           | 21°18' | 157°52' | Harbor        | 408 | 24  | 34    |
| Mexico             |        |         |               |     |     |       |
| Salina Cruz        | 16°10' | 95°12'  | Harbor        | 80  | 9   | 18    |
| Acapulco           | 16°51' | 99°55'  | Harbor        | 84  | 9   | 19    |
| Mazatlan           | 23°11' | 106°26' | Harbor        | 72  | 9   | 16    |
| La Paz             | 24°10' | 110°19' | Bay channel   | 84  | 9   | 19    |
| Guaymas            | 27°55' | 110°55' | Harbor        | 77  | 9   | 17    |
| California         |        |         |               |     |     |       |
| La Jolla           | 32°52' | 117°15' | Open coast    | 408 | 24  | 34    |
| Los Angeles        | 33°43' | 118°16' | Harbor        | 420 | 24  | 35    |
| Avila              | 35°10' | 120°44' | Open coast    | 150 | 12  | 25    |
| San Francisco      | 37°48' | 122°28' | Bay entrance  | 408 | 24  | 34    |
| Crescent City      | 41°45' | 124°12' | Harbor        | 300 | 24  | 25    |
| Oregon             |        |         |               |     |     |       |
| Astoria            | 46°13' | 123°46' | River mouth   | 372 | 24  | 31    |
| Washington         |        |         |               |     |     |       |
| Neah Bay           | 48°22' | 124°37' | Small bay     | 276 | 24  | 23    |
| Seattle            | 47°36' | 122°20' | Sound         | 408 | 24  | 34    |
| British Columbia   |        |         |               |     |     |       |
| Victoria           | 48°25' | 123°22' | Harbor        | 192 | 24  | 16    |
| Alert Bay          | 50°36' | 126°56' | Strait        | 126 | 12  | 21    |
| Alaska             |        |         |               |     |     |       |
| Ketchikan          | 55°20' | 131°38' | Narrow strait | 444 | 24  | 37    |
| Sitka              | 57°03' | 135°20' | Harbor        | 192 | 24  | 16    |
| Yakutat            | 59°33' | 139°44' | Small bay     | 216 | 24  | 18    |
| Womans Bay, Kodiak | 57°43' | 152°31' | Small bay     | 108 | 12  | 18    |

is the total number of observations and  $m$  the number of spectral estimates, was at least 15.

*Data.* Monthly mean sea levels were obtained from the U. S. Coast and Geodetic Survey, the Canadian Hydrographic Survey, and the Instituto de Geofisica UNAM, Mexico, for stations in the respective countries. The monthly means were computed from daily means of hourly heights according to  $z = 1/24(h_0 + h_1 + \dots + h_{23})$ , where  $z$  is the daily mean and  $h$  refers to the hourly value. This particular kind of averaging largely eliminates the influence of diurnal, semidiurnal, and shorter tidal components, but it does not discriminate against long-period tides. The latter are of such a small amplitude, however, that they may be neglected in the first approximation. Much more serious contamination of daily sea levels can be expected in semiencllosed bays and harbors, but the elimination of these is difficult and has not been attempted here. The accuracy of a single

hourly sea level reading is estimated to be  $\pm 1$  cm.

Monthly mean sea level atmospheric pressures were obtained from the U. S. Weather Bureau [*Monthly Weather Review*, 1872-1956], the Canadian Department of Transport, and the Servicio Meteorologico Mexicano. For the Canadian and United States stations the monthly averages are computed from five equidistant daily readings at 0000, 0600, 1200, 1800 GMT, and for Mexican stations from three daily readings at 0300, 1300, and 2100 GMT. The accuracy of a particular reading is estimated to be 0.2 mb.

Monthly mean sea surface temperatures were obtained from the U. S. Coast and Geodetic Survey [1956]. They represent, for the most part, averages of one daily reading taken at a random hour during the forenoon. The accuracy of a single reading is about 0.3°C.

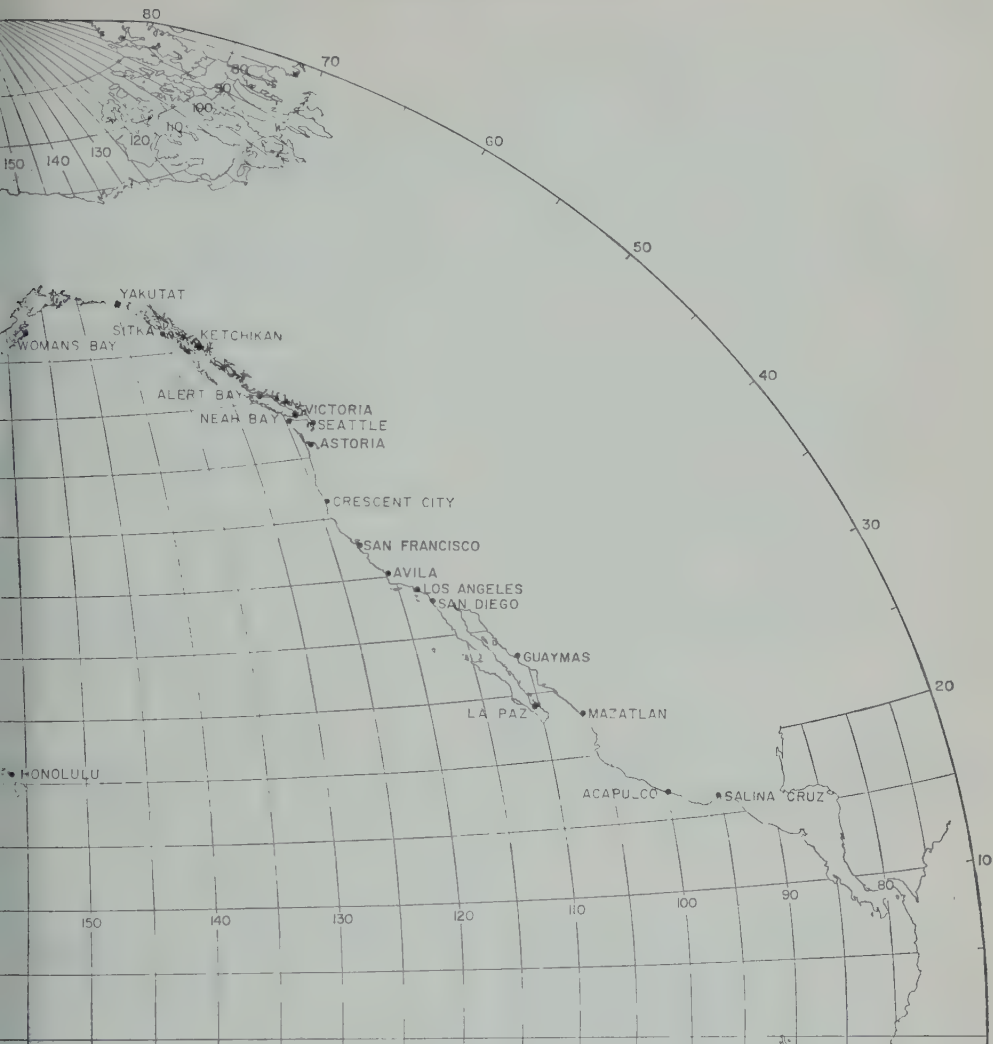


Fig. 1. Location of tide gage station used in this investigation.

records for the Columbia River were furnished by the *U. S. Geological Survey* [1958]. Averages are obtained from daily readings. The accuracy of a single determination is difficult to estimate, but should be of the order of  $\text{ft}^3/\text{sec}$ .

The observed sea level anomalies are illustrated in Figures 2a to 2d. Although the details of these records vary, a few preliminary results can be obtained from a visual inspection of the records; (a) the magnitude of the anomalies increases toward the north, (b) the maximum magnitude of the anomalies appears to be in the

neighborhood of 30 cm, (c) there are no obvious periodicities discernible, (d) the anomalies are similar over distances of several hundred kilometers, and (e) for the majority of stations the mean of the record does not change appreciably.

The probability distribution of sea level, atmospheric pressure, and sea surface temperature anomalies is shown in Figure 3. It is seen that the distribution is approximately normal, except at the extremes.

The standard deviations and extreme ranges of the anomalies for each of the stations are given in Table 2. For sea level and atmospheric

pressure anomalies these are seen to increase from south to north, the increase being by a factor of roughly 2 for the former and about 6 for the latter. For sea surface temperature anomalies there is no dependence upon latitude; the anomalies appear, however, to be larger for stations located in shallow bays and harbors than for those situated in straits and sounds, suggesting that tidal mixing may be important in reducing the amplitude of the temperature variation.

The standard deviation of the sea level anomalies can, in many cases, be reduced by eliminating the static effect of atmospheric pressure from the sea level records. If this is done (noting that an excess of pressure of 1 mb decreases the sea level by 1 cm) the resulting standard deviation of the 'corrected' sea level is seen in the last column of Table 2 to depend very little upon latitude. However, a significant difference between the observed and corrected sea level anomalies exists in the northern part of the region only.

*Computational procedures.* Let  $z(t)$  and  $p(t)$  denote anomalies of sea level and atmospheric pressure. Then the autocorrelations and cross correlations are defined by

$$\phi_{zz}(\tau) = \langle z(t)z(t + \tau) \rangle, \tag{1}$$

$$\phi_{zp}(\tau) = \langle z(t)p(t + \tau) \rangle$$

respectively, provided the records are stationary. Here  $\langle \rangle$  denotes the time average over the entire record length, and  $\tau$  the time lag in months.

The power spectral densities and cross-power spectral densities are obtained from a Fourier transform of the correlation functions:

$$E_{zz}(\omega) = \frac{2}{\pi} \int_0^\infty \phi_{zz}(\tau) \cos \omega \tau d\tau, \tag{2}$$

$$E_{zp}(\omega) = \frac{1}{\pi} \int_{-\infty}^\infty \phi_{zp}(\tau) e^{i\omega\tau} d\tau$$

Practical methods for evaluating the power spectra are given by *Blackman and Tukey* [1958] and *Munk, Snodgrass, and Tucker* [1959].

The coherence between  $z(t)$  and  $p(t)$  is defined by

$$R_{zp}(\omega) = \frac{E_{zp}(\omega)}{[E_{zz}(\omega)E_{pp}(\omega)]^{1/2}} \tag{3}$$

and must lie between the values 0 and 1.

The phase lag of  $z(t)$  over  $p(t)$  is defined by

$$\pi_{zp}(\omega) = \arctan \frac{E_{zp} - E_{pz}}{E_{zp} + E_{pz}}$$

provided that for a positive  $E_{zp} - E_{pz}$   $0 \leq \pi \leq 180$  and for a negative numerator  $180 \leq \pi \leq 360$ .

It should be noted that the autocorrelation and therefore also the spectral densities, depend upon the amount of smoothing or averaging performed upon the original data. In the case of a Markov process, the autocorrelation is improved by an increased amount of averaging [*Munk, Snodgrass, and Tucker*, 1960]. Thus a valid comparison

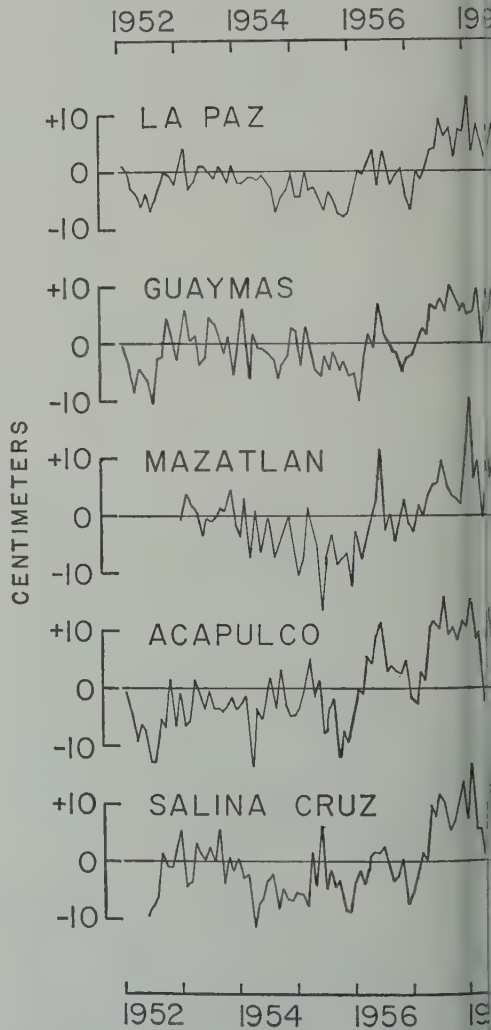


Fig. 2a. Sea level anomalies along the coast of Mexico.

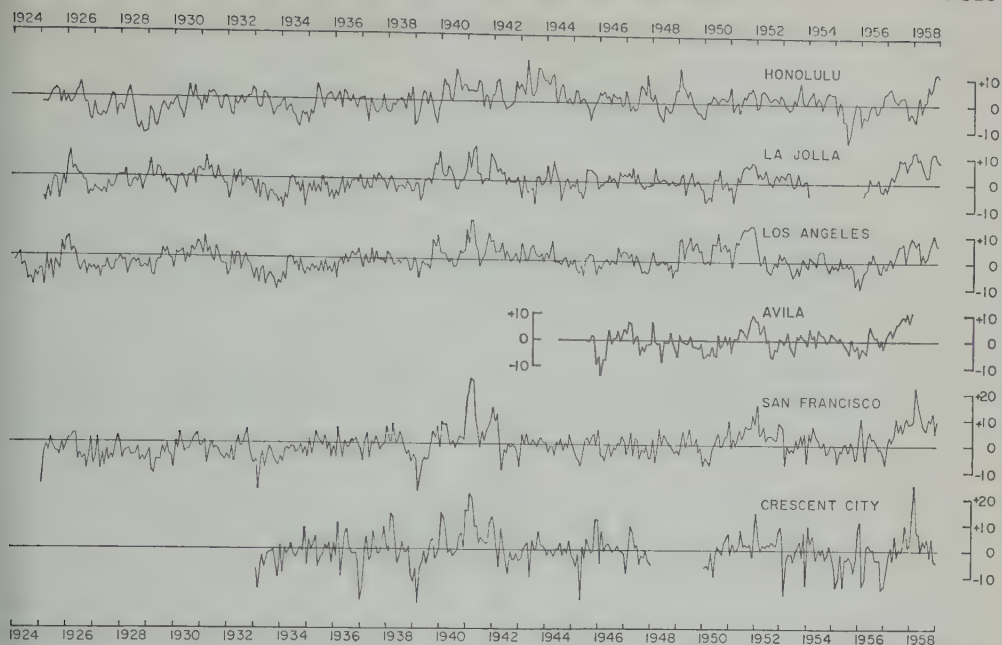


Fig. 2b. Sea level anomalies along the coast of California and Hawaii.

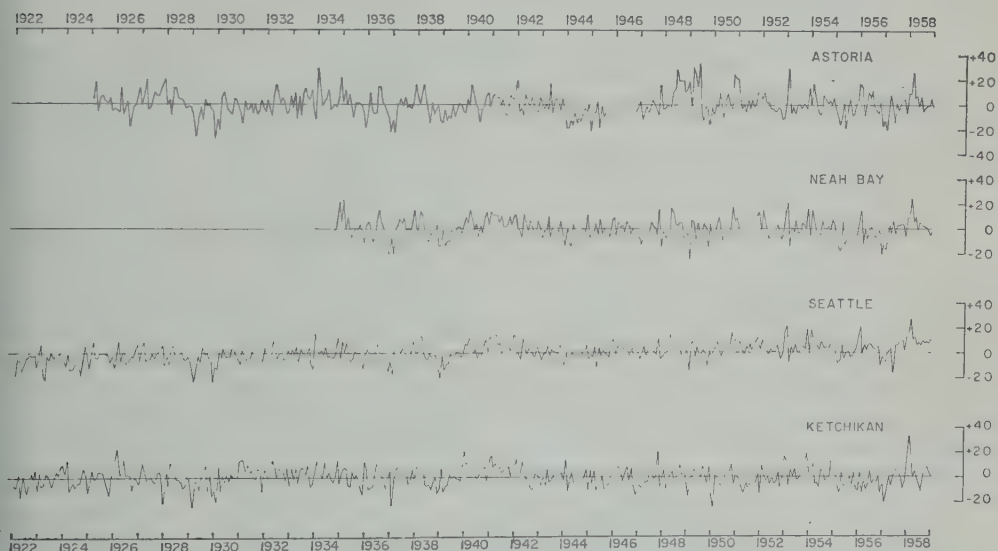


Fig. 2c. Sea level anomalies along the coasts of Oregon, Washington, and Alaska.

on between autocorrelations or spectral ties can be made only if the data are averaged in the same way. In the present case the averaging was the same for each variable at all stations, though different for different variables.

The discussion below is therefore limited to differences between stations.

The 95 per cent confidence limits for the power spectra coherence and phase are given in Table 3. They are based on the assumption



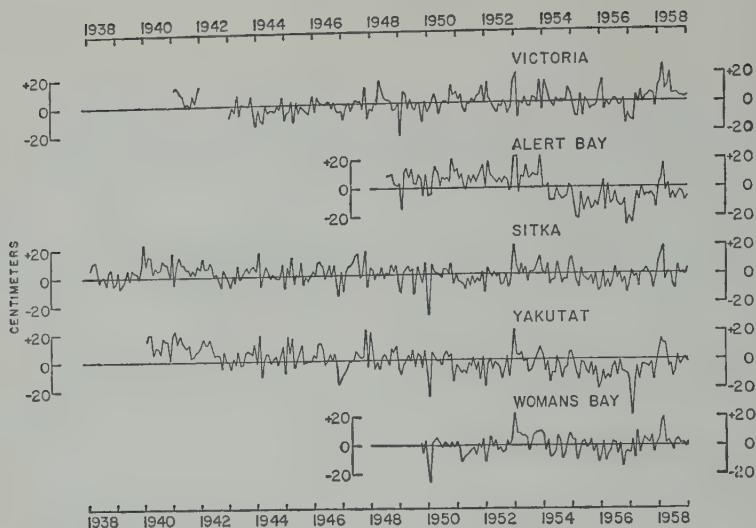


Fig. 2d. Sea level anomalies along the coast of British Columbia and Alaska.

that the anomalies are normally distributed and also possess a joint normal distribution. The confidence limits for the coherence and phase computed by Goodman [1957], have recently been questioned by Tick (personal communication). For the autocorrelation Laning and Battin [1956] give the following inequality from which the standard deviation  $\sigma$  of the observed autocorrelation from the true autocorrelation of an infinite record can be obtained:

$$\sigma \leq \left[ \frac{4}{n} \int_0^\infty \phi^2(t) dt \right]^{1/2} \quad (5)$$

where  $n$  denotes the total record length.

*Response between sea level and atmospheric pressure anomalies.* The time required for the sea surface to come into equilibrium with the field of atmospheric pressure has received considerable attention recently. In the case of an infinite and homogeneous ocean, Sarkisyan [1957a, b] computed the equilibrium time to be of the order of 1 to 2 months. For an inhomogeneous ocean Veronis and Stommel [1956] found that at least 100 years would be required to reach equilibrium. It appears, therefore, that for periods shorter than a month the sea surface does not adjust itself completely to the field of atmospheric pressure, regardless of whether an isotropic or baroclinic ocean is considered.

Once the equilibrium is established the de-

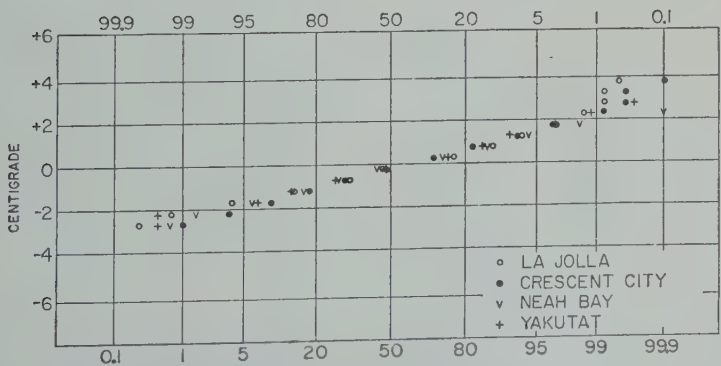
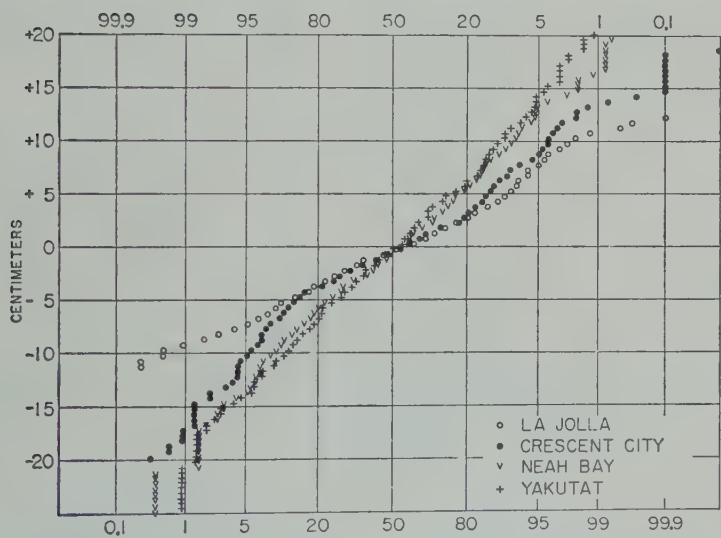
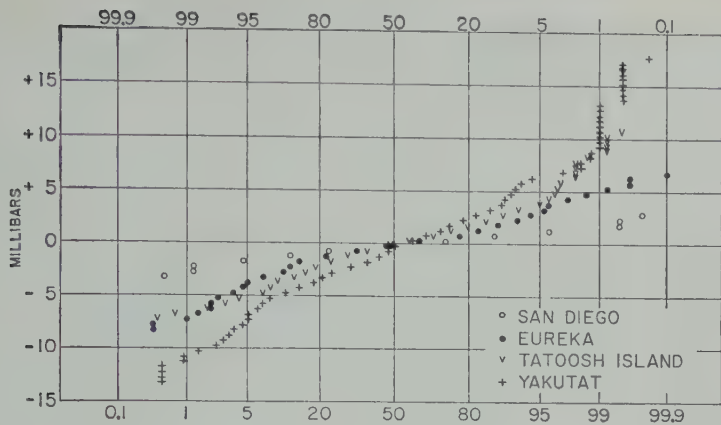
parture of the sea surface from equilibrium can be computed from the Bernoulli equation

$$\eta = -p_0/\rho g - \mathbf{v}^2/2g$$

where  $\mathbf{v}$  is the vector velocity,  $\rho$  the density,  $g$  the acceleration of gravity,  $p_0$  the atmospheric pressure, and  $\eta$  the height of sea level above equilibrium surface. This equation has been derived on the assumption that the flow is irrotational and takes place on a resting earth. If rotation of the earth is also considered, an integrated Coriolis term must be added.

The first term represents the effect of atmospheric pressure upon sea level; in particular it shows that, in the absence of motion ( $\mathbf{v} = 0$ ), the sea level is depressed by approximately 1 cm for every millibar of pressure increase. The second term takes into account the dynamic effect, and it can be seen that if the atmospheric pressure is constant the sea level is lowered 1 cm for every increase in current of about 45 cm/sec. Such current changes are seldom observed, however, and the dynamic term, being much smaller than the static term, is generally neglected.

On a rotating earth, the sea level is further influenced by the Coriolis force; a rough estimate of the magnitude of this effect can be obtained by considering a current flowing northward along the west coast of North America.



3. Probability distribution of atmospheric pressure, sea level and sea surface temperature anomalies. Dots refer to 0.5 unit class intervals.

TABLE 2. Standard Deviations ( $\sigma$ ) and Extreme Ranges ( $r$ ) of Anomalies of Sea Level ( $z$ ) Atmospheric Pressure ( $p$ ), and Sea Surface Temperature ( $\theta$ )

| Station                   | $\sigma_z$<br>cm | $\sigma_p$<br>mb | $\sigma_\theta$<br>°C | $r_z$<br>cm | $r_p$<br>mb | $r_\theta$<br>°C | ( $\sigma_z$ ) cor<br>cm |
|---------------------------|------------------|------------------|-----------------------|-------------|-------------|------------------|--------------------------|
| Hawaii                    |                  |                  |                       |             |             |                  |                          |
| Honolulu                  | 4.6              | 1.2              | 0.5                   | 30          | 8           | 3                | 4.6                      |
| Mexico                    |                  |                  |                       |             |             |                  |                          |
| Salina Cruz               | 5.7              | 0.8              | 1.4                   | 28          | 4           | 5                | 5.5                      |
| Acapulco                  | 6.5              | 0.8              | 0.9                   | 29          | 4           | 4                | 6.2                      |
| Mazatlan                  | 6.5              | 0.7              | 1.7                   | 32          | 3           | 6                | 6.4                      |
| La Paz                    | 4.3              | 0.7              | 0.8                   | 21          | 4           | 4                | 4.2                      |
| Guaymas                   | 4.6              | 0.6              | 1.1                   | 20          | 4           | 4                | 4.6                      |
| California                |                  |                  |                       |             |             |                  |                          |
| La Jolla-San Diego*       | 4.2              | 0.9              | 1.0                   | 24          | 8           | 6                | 4.0                      |
| Los Angeles               | 4.6              | 0.9              | 1.1                   | 27          | 6           | 7                | 4.2                      |
| Avila-Santa Maria         | 4.1              | 0.9              | 0.6                   | 25          | 5           | 3                | 4.0                      |
| San Francisco             | 5.5              | 1.5              | 0.8                   | 43          | 12          | 5                | 5.0                      |
| Crescent City-Eureka      | 6.1              | 2.1              | 1.1                   | 45          | 14          | 6                | 5.0                      |
| Oregon                    |                  |                  |                       |             |             |                  |                          |
| Astoria-North Head        | 9.2              | 2.6              | 1.0                   | 64          | 15          | 7                | 7.9                      |
| Washington                |                  |                  |                       |             |             |                  |                          |
| Neah Bay-Tatoosh Island   | 7.7              | 2.7              | 1.0                   | 49          | 18          | 5                | 5.6                      |
| Seattle                   | 6.7              | 2.4              | 0.8                   | 45          | 16          | 5                | 5.2                      |
| British Columbia          |                  |                  |                       |             |             |                  |                          |
| Victoria†                 | 6.7              | 2.7              | 0.5                   | 47          | 18          | 3                | 4.9                      |
| Alert Bay‡ Port Hardy     | 10.0             | 3.0              | 0.5                   | 48          | 14          | 3                | 8.8                      |
| Alaska                    |                  |                  |                       |             |             |                  |                          |
| Ketchikan-Annette         | 7.6              | 3.5              | 0.8                   | 57          | 23          | 4                | 5.6                      |
| Sitka-Juneau              | 6.5              | 3.9              | 0.8                   | 48          | 24          | 4                | 4.4                      |
| Yakutat                   | 8.5              | 4.1              | 1.0                   | 60          | 31          | 6                | 6.8                      |
| Womans Bay, Kodiak Island | 6.8              | 5.3              | 0.9                   | 49          | 36          | 4                | 5.6                      |

\* Where two names occur, the first refers to tide gage stations, the second to meteorological ones.

† Sea surface temperature measured at Race Rocks, B.C.

‡ Sea surface temperature measured at Pine Island, B.C.

with an average width of 100 km and an average velocity of 10 cm/sec, assumed to be constant from top to bottom. The rise in sea level due solely to the Coriolis force would be about 10 cm. It is thus seen that relatively small changes in the average current speed can lead to appreciable changes in sea level. Fortunately, the known changes in the speed of the average current along the west coast of North America over a time period of a month are not very large, mostly within 20 cm/sec, and therefore excessively large sea level variations due to the Coriolis force alone are not to be expected.

*Autocorrelation and power spectral density.* The autocorrelation of sea level anomalies is shown in Figure 4. The dots refer to the observed anomalies, the crosses to those from which the atmospheric pressure effect has been

eliminated. The dashed line refers to the standard deviation of the computed autocorrelation from the true one of an infinite record. It is seen that the persistence of the anomalies adjusted for the atmospheric pressure effect is improved over the persistence of the unadjusted anomalies, particularly at latitudes north of San Francisco. The autocorrelation differs significantly from zero for lag times between 0 and 4 to 16 months, depending upon the individual station but not upon latitude. Most of the autocorrelations suggest an exponential decay of the sea level anomaly with time, at least for the first few lags.

The spectral density of the anomalies for the frequency range between 0 and 6 cycles per year (cpy) is shown in Figure 5, where the dots refer, as previously, to the observed anomalies

TABLE 3. Conference Limits (95 %) for the Power Spectra ( $E$ ), the Coherence ( $R$ ), and the Phase ( $\pi$ )

| Station                      | $E$ ,<br>cm <sup>2</sup> /cpy | $R$  | $\pi$ ,<br>deg. |
|------------------------------|-------------------------------|------|-----------------|
| Hawaii                       |                               |      |                 |
| Honolulu                     | 0.66-1.71                     | 0.34 | 41              |
| Puerto Rico                  |                               |      |                 |
| Salina Cruz                  | 0.57-2.19                     | 0.47 | 38              |
| Acapulco                     | 0.58-2.13                     | 0.46 | 38              |
| Mazatlan                     | 0.56-2.31                     | 0.50 | 37              |
| La Paz                       | 0.58-2.13                     | 0.46 | 38              |
| Guaymas                      | 0.56-2.26                     | 0.48 | 38              |
| California                   |                               |      |                 |
| La Jolla                     | 0.66-1.71                     | 0.34 | 41              |
| Los Angeles                  | 0.66-1.69                     | 0.34 | 41              |
| Avila                        | 0.62-1.90                     | 0.40 | 40              |
| San Francisco                | 0.66-1.71                     | 0.34 | 41              |
| Crescent City                | 0.62-1.90                     | 0.40 | 40              |
| Oregon                       |                               |      |                 |
| Astoria                      | 0.64-1.78                     | 0.36 | 41              |
| Washington                   |                               |      |                 |
| Neah Bay                     | 0.60-1.97                     | 0.42 | 39              |
| Seattle                      | 0.66-1.71                     | 0.34 | 41              |
| British Columbia             |                               |      |                 |
| Victoria                     | 0.56-2.31                     | 0.50 | 37              |
| Alert Bay                    | 0.59-2.04                     | 0.44 | 39              |
| Alaska                       |                               |      |                 |
| Ketchikan                    | 0.66-1.67                     | 0.33 | 41              |
| Sitka                        | 0.56-2.31                     | 0.50 | 37              |
| Yakutat                      | 0.57-2.19                     | 0.47 | 38              |
| Womans Bay,<br>Kodiak Island | 0.57-2.19                     | 0.47 | 38              |

crosses to those adjusted for the atmospheric pressure effect. The main feature in the spectra is that they contain most of their power in the low-frequency part and almost nothing in higher frequencies. As can be expected from autocorrelations, the spectrum of the anomalies from which the atmospheric pressure effect has been eliminated contains less power in the spectrum of the observed anomalies, indicating that a substantial part of the sea level variations north of San Francisco is really due to variations in atmospheric pressure. It is also worthy that there are no significant peaks in the spectra, which emphasizes the absence of periodicities, a fact already mentioned by *Brich and Munk* [1959].

The autocorrelations and power spectral densities of atmospheric pressure anomalies (not shown here) indicate a persistence of less than one month and a distribution of power almost in-

dependent of frequency. For sea surface temperature anomalies the autocorrelations and spectral densities are similar to those in sea level anomalies; the persistence varies between 4 and 6 months and is independent of latitude; most of the power is found at frequencies between 0 and 2 cpy.

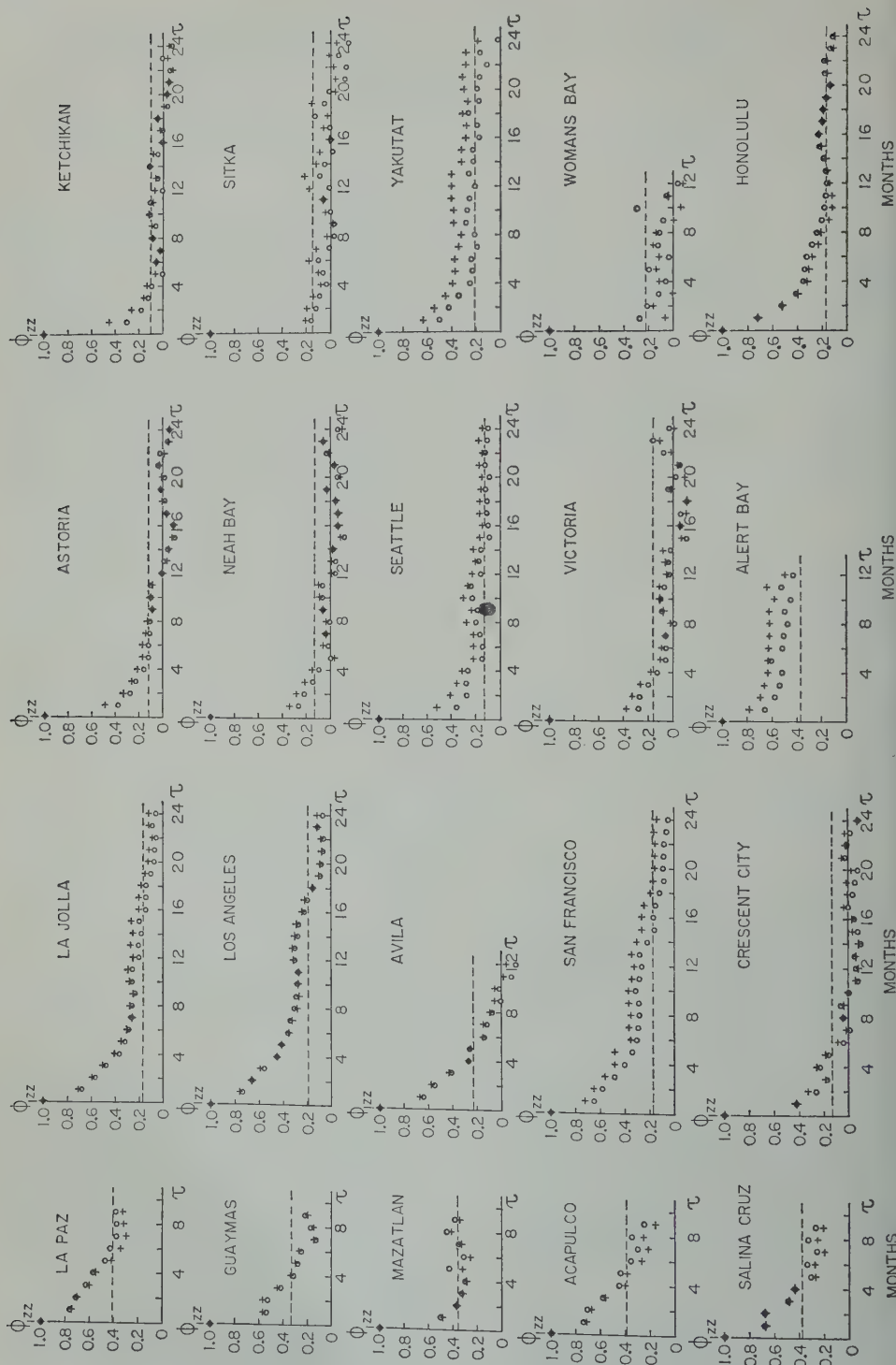
*Coherence and phase between sea level and atmospheric pressure anomalies.* The coherence spectrum is interesting from two points of view. It shows that the coherence along the coast becomes better toward the north and that at most stations the coherence depends very little upon frequency (Figs. 6 and 7). The latter may be interpreted to mean that sea level adjusts itself rather quickly to variations in atmospheric pressure, certainly in less than 2 months. The highest values of the coherence are about 0.8, indicating that about 64 per cent of the sea level anomalies can be accounted for in terms of atmospheric pressure anomalies. In the southern part of the region the coherence varies around 0.4, showing that only 16 per cent of the sea level anomalies are related to atmospheric pressure variations. The phase between these variables is 180° and, for all except the Mexican stations, varies very little with frequency. This indicates that the sea level responds to atmospheric pressure like an 'inverted barometer,' for most stations investigated here. Why the phase scatters so much at the Mexican stations is difficult to explain; it is possible that the nonseasonal atmospheric pressure variations are small and therefore have no effect upon sea level.

One of the most interesting and yet unexplained results is the ratio of the energy of sea level anomalies to the energy of atmospheric pressure anomalies as a function of frequency (Table 4).

It is seen that that ratio decreases sharply with increasing frequency and that the ratios at low frequencies are much smaller in the northern than in the southern part of the region. This suggests that something other than atmospheric pressure must also influence sea level. If it were not so, the ratio should be equal to 1 at all frequencies.

*Coherence and phase between sea level and sea surface temperature anomalies.* The coherence spectra (Fig. 8) do not show a relation to latitude, but seem rather to depend





○ OBSERVED SEA LEVEL ANOMALIES

+ CORRECTED SEA LEVEL ANOMALIES

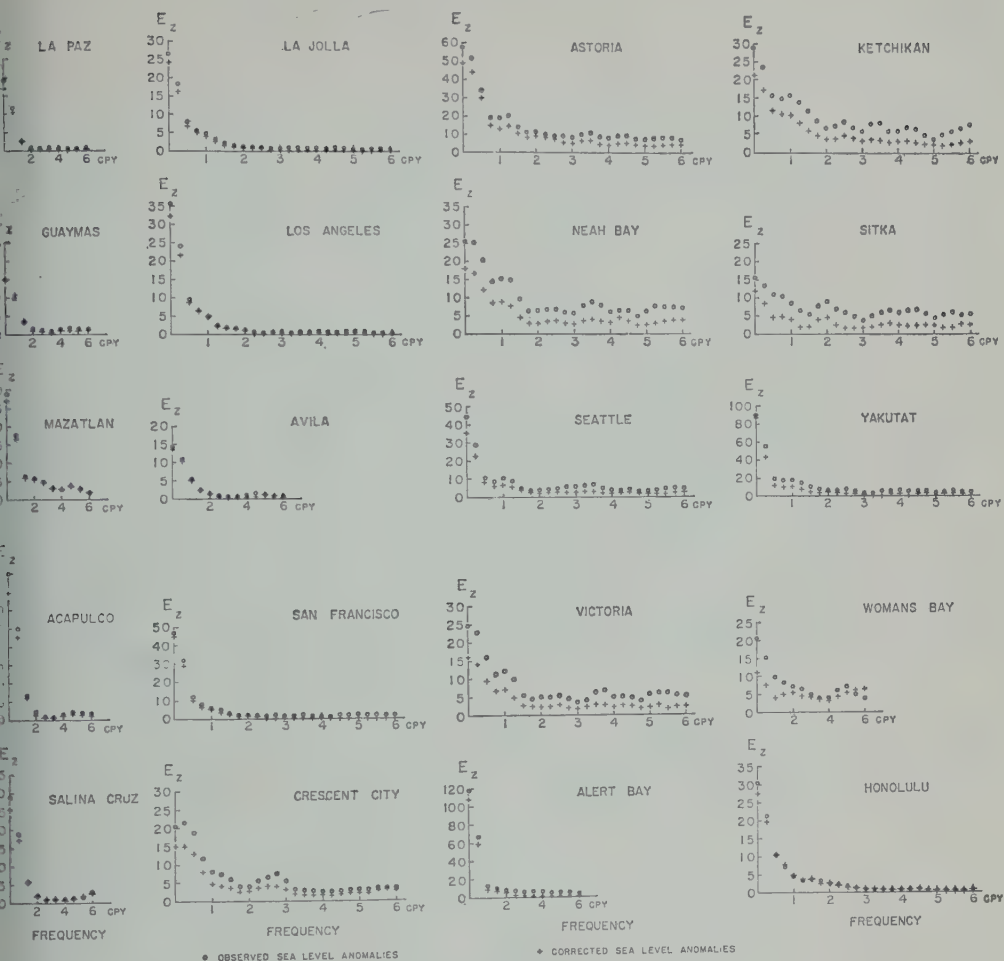


Fig. 5. Energy of sea level anomalies.

the particular environment in which the gage station is located. A relatively good coherence is found in most shallow bays and harbors, such as Los Angeles, Neah Bay, and Yakutat, whereas the coherence is poor in such places as river mouths (Astoria), and straits and sounds (Victoria, Seattle). Even along the open coast (La Jolla) the coherence is only moderate. In many cases the coherence is seen to depend on frequency, but the relation is not clear cut. On an average, less than 20 per cent of the sea level anomalies appear to be due to anomalies in surface temperature (Fig. 9). The phase between these anomalies is zero where the coherence is high and indicates a direct relation between the height of sea level and local warm-

ing. In places where tidal mixing is considerable, such as in straits and sounds, there is much scatter in the phase.

*Coherence and phase between sea level and wind anomalies.* The coherence between anomalies in the south component of the geostrophic wind and the sea level has been investigated for three places along the west coast of the United States in order to find out whether the Coriolis force has any effect upon sea level (Fig. 10). It is seen that the coherence is rather poor off Southern California, and somewhat better near Cape Mendocino and Cape Flattery. But even in the best cases not more than about one-third of the sea level variations can be related to variations in the wind. The phase between

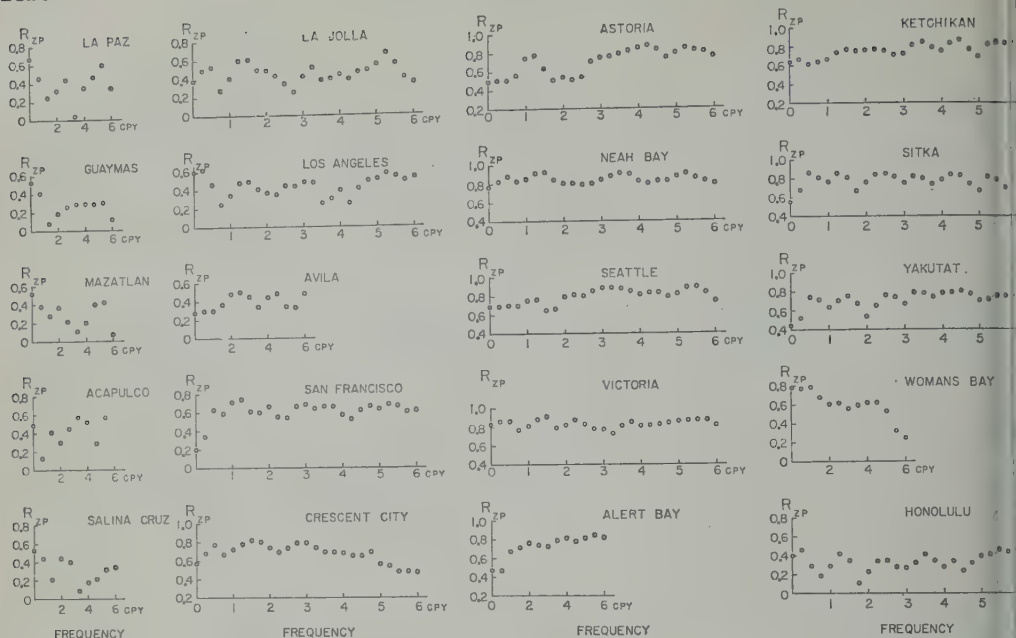


Fig. 6. Coherence between sea level and atmospheric pressure anomalies.

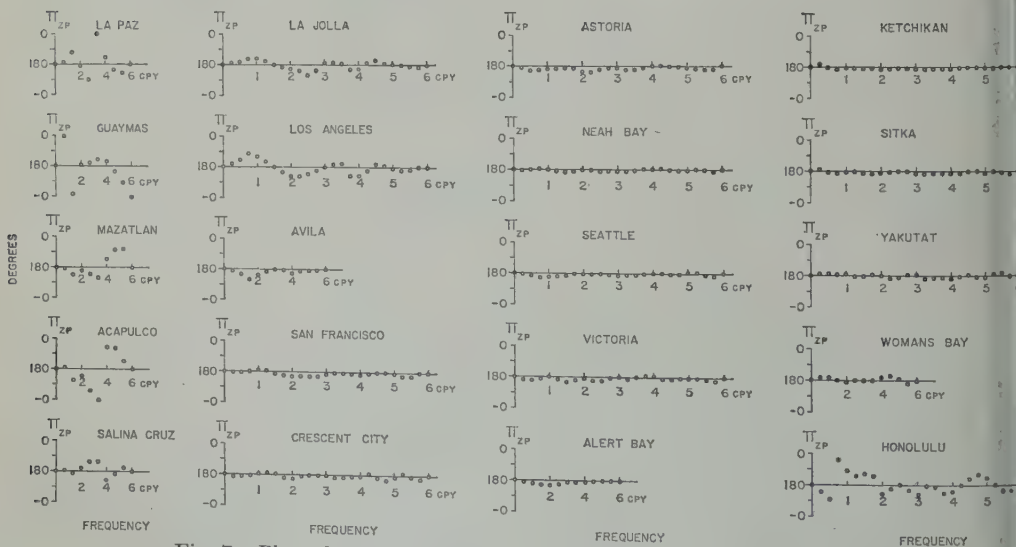


Fig. 7. Phase between sea level and atmospheric pressure anomalies.

these variables is  $0^\circ$  in the northern part of the region, indicating that a south wind leads to high sea levels, as can be expected from the deviating force of the earth's rotation. It is noteworthy that the phase in the north is almost constant with frequency. Off Southern California the phase is also zero where the coherence

is high, but since the winds there are more northerly this indicates that the sea level is lowered when the winds are strong. It may be concluded that the Coriolis force has a significant but recognizable effect upon sea level along the west coast of the United States.

*Coherence and phase between sea level*

TABLE 4. The Ratio  $E_s/E_p$  as a Function of Frequency

| Frequency, cpy | Tide Station |          |               |          |           |
|----------------|--------------|----------|---------------|----------|-----------|
|                | Honolulu     | La Jolla | San Francisco | Neah Bay | Ketchikan |
| 0              | 37.58        | 74.07    | 44.79         | 20.42    | 15.54     |
| 1              | 16.48        | 20.72    | 12.25         | 11.41    | 7.38      |
| 2              | 11.47        | 9.91     | 6.04          | 4.64     | 3.56      |
| 3              | 7.12         | 5.00     | 4.70          | 5.17     | 2.60      |
| 4              | 5.59         | 6.84     | 4.77          | 6.96     | 3.32      |
| 5              | 3.10         | 5.39     | 4.99          | 4.44     | 2.64      |
| 6              | 3.28         | 5.54     | 5.60          | 6.14     | 3.06      |

ff anomalies. The considerable influence of the discharge of the Columbia River upon the sea level at Astoria, Oregon, is illustrated by the relatively high coherence, particularly at low frequencies (Fig. 11). Depending upon frequency, between 15 per cent and 50 per cent of sea level variations can be related to variations in discharge. The phase scatters around zero showing that the variables are directly related. This indicates that tide gage stations in river mouths are not representative and cannot be used in studies of sea level variations on a larger than local scale.

*Statistical prediction of sea level anomalies.* The predicted sea level anomaly  $z^*$  at any time month  $\alpha$  be a linear combination of the present and past anomalies of sea level, atmospheric pressure, and sea surface temperature.

$$z^* + \alpha) = \sum_{i=0}^n [a_i z(t-i) + \lambda_{sp} b_i p(t-i) + \lambda_{ss} c_i \theta(t-i)] \quad (7)$$

where  $a_i$ ,  $b_i$ ,  $c_i$  are weighting coefficients to be determined by the method of least squares and  $n$  is the number of past months taken into account, and where the constants

$$\lambda_{sp} = \left[ \frac{\langle z^2 \rangle}{\langle p^2 \rangle} \right]^{1/2} \quad \lambda_{ss} = \left[ \frac{\langle z^2 \rangle}{\langle \theta^2 \rangle} \right]^{1/2}$$

are introduced to keep the weighting coefficients dimensionless. The success or failure of the prediction can be described in terms of the mean-square error  $\epsilon^2$ , defined by

$$\epsilon^2 = \frac{\langle [z^* - z]^2 \rangle}{\langle z^2 \rangle} \quad (8)$$

Methods for the computation of the weighting coefficients and the mean-square error are discussed elsewhere [Roden and Groves, 1960] and will not be repeated.

Here only two simple cases will be considered. The first case concerns the estimation of the present sea level anomaly from the present anomalies of atmospheric pressure and sea surface temperature. Instead of equation 6 we have

TABLE 5. Coefficients and Mean-Square Error for an Estimate of the Present Sea Level Anomaly from the Present Anomalies of Atmospheric Pressure and Sea Surface Temperature

|                  | $\lambda_{sp}$<br>cm/mb | $\lambda_{ss}$<br>cm/°C | $b_0$ | $c_0$ | $\epsilon^2$ |
|------------------|-------------------------|-------------------------|-------|-------|--------------|
| Hawaii           |                         |                         |       |       |              |
| Honolulu         | 3.8                     | 9.2                     | 0.07  | 0.21  | 0.98         |
| Mexico           |                         |                         |       |       |              |
| Acapulco         | 8.1                     | 7.2                     | -0.28 | 0.61  | 0.58         |
| La Paz           | 6.1                     | 5.4                     | -0.33 | 0.36  | 0.76         |
| California       |                         |                         |       |       |              |
| La Jolla         | 4.7                     | 4.2                     | -0.27 | 0.65  | 0.63         |
| Los Angeles      | 5.1                     | 4.2                     | -0.20 | 0.77  | 0.51         |
| Avila            | 4.6                     | 6.8                     | -0.24 | 0.49  | 0.73         |
| San Francisco    | 3.7                     | 6.9                     | -0.44 | 0.39  | 0.67         |
| Crescent City    | 2.9                     | 5.5                     | -0.62 | 0.44  | 0.44         |
| Oregon           |                         |                         |       |       |              |
| Astoria          | 3.5                     | 9.2                     | -0.54 | 0.02  | 0.71         |
| Washington       |                         |                         |       |       |              |
| Neah Bay         | 2.8                     | 7.7                     | -0.74 | 0.67  | 0.06         |
| Seattle          | 2.8                     | 8.4                     | -0.73 | 0.04  | 0.47         |
| British Columbia |                         |                         |       |       |              |
| Victoria         | 2.5                     | 11.2                    | -0.78 | 0.27  | 0.31         |
| Alert Bay        | 3.3                     | 20.0                    | -0.56 | 0.16  | 0.66         |
| Alaska           |                         |                         |       |       |              |
| Ketchikan        | 2.2                     | 9.5                     | -0.69 | 0.28  | 0.44         |
| Sitka            | 1.7                     | 8.1                     | -0.72 | 0.38  | 0.34         |
| Yakutat          | 2.1                     | 8.5                     | -0.58 | 0.54  | 0.39         |



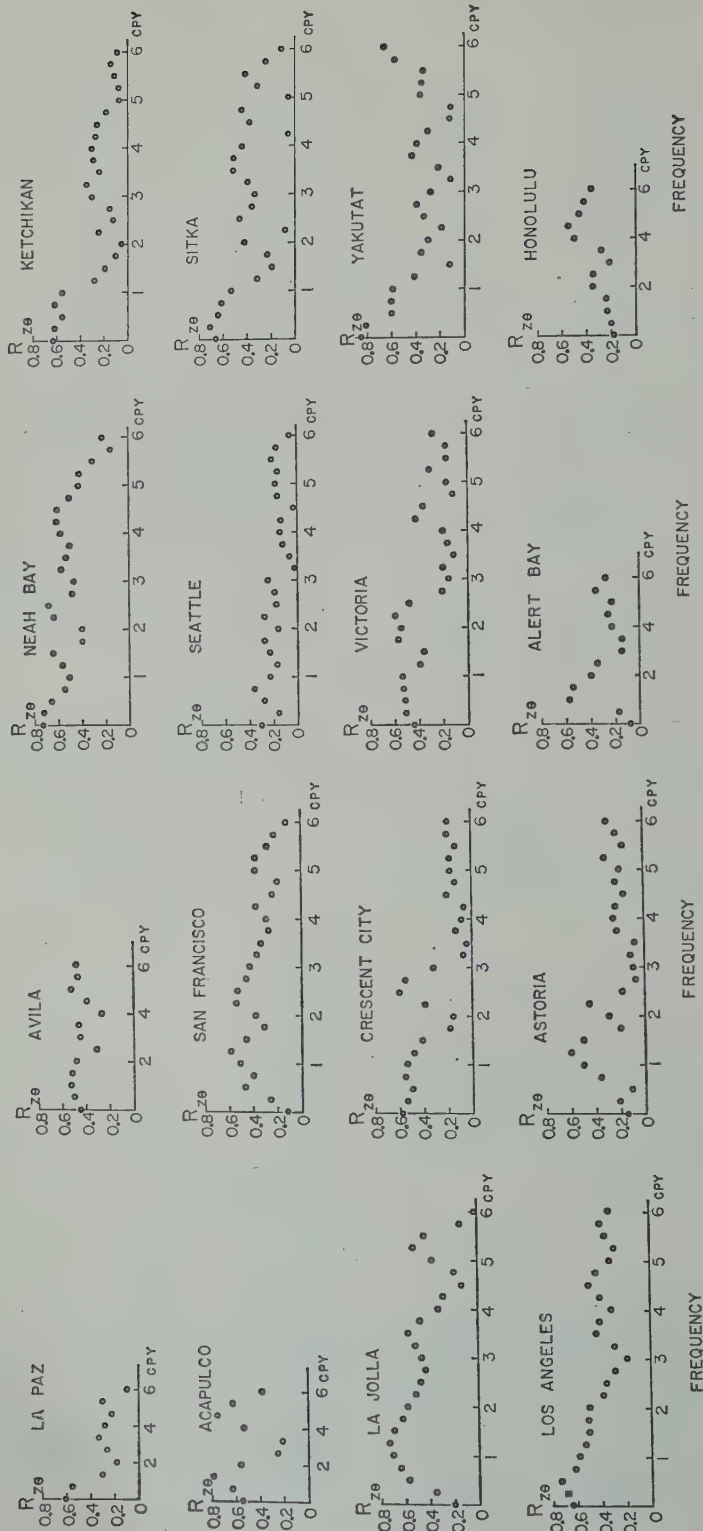


Fig. 8. Coherence between sea level and sea surface temperature anomalies.

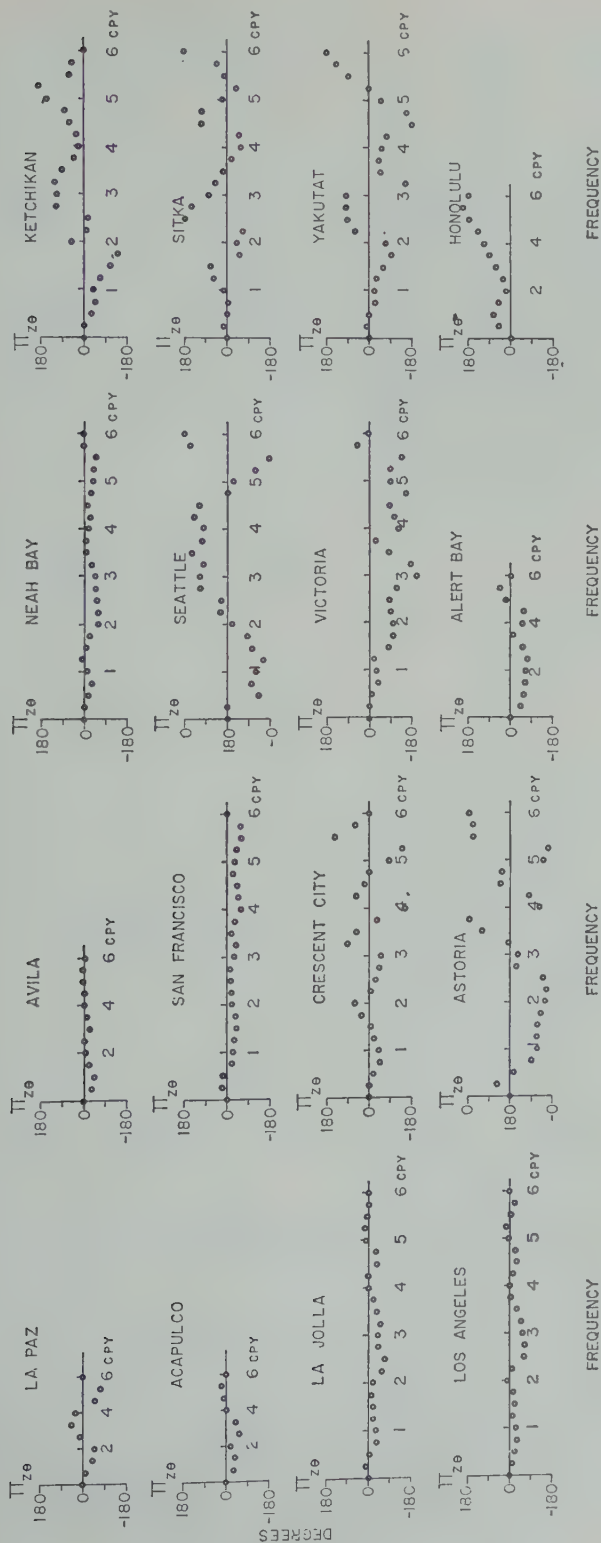


Fig. 9. Phase between sea level and sea surface temperature anomalies.

$$z^*(0) = \lambda_{zp}b_0p(0) + \lambda_{z\theta}c_0\theta(0) \tag{9}$$

The coefficients and the mean-square error are given in Table 5. It is seen that for stations north of San Francisco a fairly good estimate of the sea level anomaly can be made; on an average more than 50 per cent can be related to anomalies in atmospheric pressure and sea surface temperature. In the case of Neah Bay, Washington, this percentage is as high as 94 per cent. For stations in Mexico and Southern California the estimates are less good, and for Honolulu, Hawaii, the method fails altogether.

The second case concerns the prediction of the sea level anomaly a few months into the future from the present anomaly of sea level, atmospheric pressure, and sea surface temperature. We have the equation

$$z^*(\alpha) = a(\alpha)z(0) + \lambda_{zp}b(\alpha)p(0) + \lambda_{z\theta}c(\alpha)\theta(0) \tag{10}$$

The coefficients and the mean-square error are shown in Table 6. It is seen that the present sea level anomaly has the greatest weight, and that the other coefficients are relatively small, indicating that the prediction depends almost entirely upon the persistence of the sea level anomalies. Since the persistence is indicated by a Markov-type autocorrelation, the future value depends only upon the present and is independent of the past. Thus the inclusion of more present values of sea level anomalies in our prediction scheme would not lead to better results than we have already obtained. We must therefore conclude that prediction of sea level anomalies by the statistical method does not appear very promising.

*Acknowledgments.* I have benefited considerably from many fruitful discussions with Walter H. Munk, Maurice Blackburn, Gordon W. Grover, John D. Isaacs, and Joseph L. Reid, Jr., read the manuscript and offered many helpful suggestions.

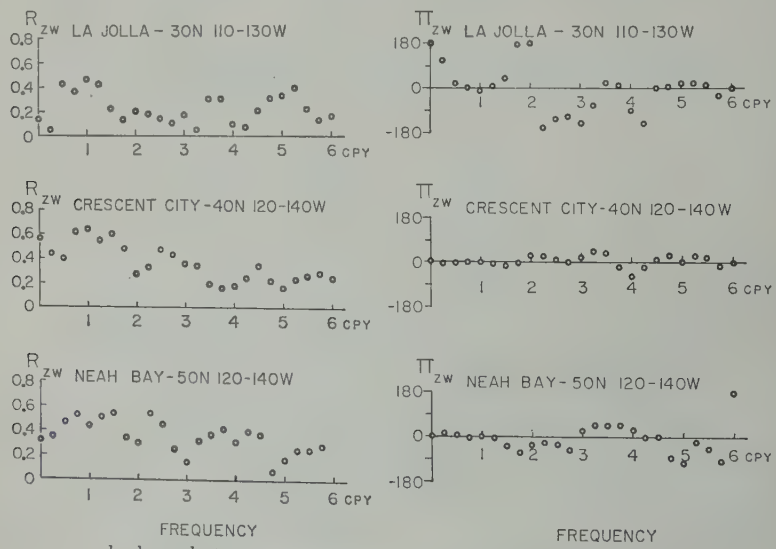


Fig. 10. Coherence and phase between anomalies of sea level and the south component of the atmospheric wind at the latitude and longitude stated.

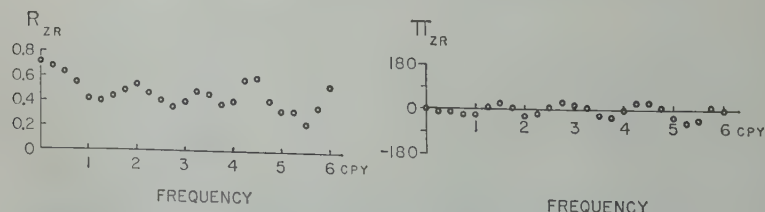


Fig. 11. Coherence and phase between sea level and runoff anomalies at Astoria, Oregon.

TABLE 6. Prediction Coefficients and Mean-Square Error for Prediction from Present Anomalies of Sea Level, Atmospheric Pressure, and Sea Surface Temperature

|               | $\lambda_{z\rho}$<br>cm/mb | $\lambda_{z\theta}$<br>cm/°C | $a$                                | $b$   | $c$   | $\epsilon^2$ | $a$                                 | $b$   | $c$   | $\epsilon^2$ |
|---------------|----------------------------|------------------------------|------------------------------------|-------|-------|--------------|-------------------------------------|-------|-------|--------------|
|               |                            |                              | $(\alpha = 1 \text{ month ahead})$ |       |       |              | $(\alpha = 2 \text{ months ahead})$ |       |       |              |
| San Francisco | 3.8                        | 9.2                          | 0.63                               | 0.02  | 0.18  | 0.57         | 0.42                                | 0.01  | 0.09  | 0.82         |
| San Francisco | 8.1                        | 7.2                          | 0.78                               | -0.08 | -0.17 | 0.44         | 0.78                                | -0.20 | -0.25 | 0.40         |
| San Francisco | 6.1                        | 5.4                          | 0.77                               | 0.04  | 0.00  | 0.43         | 0.76                                | 0.04  | -0.07 | 0.48         |
| San Francisco | 4.7                        | 4.2                          | 0.73                               | 0.09  | -0.03 | 0.52         | 0.64                                | 0.06  | -0.09 | 0.66         |
| San Francisco | 5.1                        | 4.2                          | 0.77                               | 0.11  | -0.02 | 0.46         | 0.71                                | 0.11  | -0.05 | 0.56         |
| San Francisco | 4.6                        | 6.8                          | 0.72                               | 0.05  | -0.12 | 0.53         | 0.64                                | 0.05  | -0.19 | 0.68         |
| San Francisco | 3.7                        | 6.9                          | 0.75                               | 0.18  | -0.04 | 0.55         | 0.69                                | 0.07  | -0.04 | 0.61         |
| San Francisco | 2.9                        | 5.5                          | 0.37                               | -0.02 | 0.10  | 0.81         | 0.32                                | 0.10  | 0.06  | 0.92         |
| San Francisco | 3.5                        | 9.2                          | 0.47                               | -0.10 | -0.11 | 0.80         | 0.34                                | 0.10  | -0.12 | 0.90         |
| San Francisco | 2.8                        | 7.7                          | 0.41                               | 0.27  | 0.00  | 0.93         | 0.36                                | 0.18  | 0.05  | 0.95         |
| San Francisco | 2.8                        | 8.4                          | 0.58                               | 0.30  | 0.02  | 0.83         | 0.45                                | 0.21  | 0.08  | 0.90         |
| San Francisco | 2.5                        | 11.2                         | 0.36                               | 0.12  | 0.06  | 0.91         | 0.27                                | 0.00  | 0.09  | 0.92         |
| San Francisco | 3.3                        | 20.0                         | 0.79                               | 0.40  | -0.09 | 0.56         | 0.73                                | 0.27  | -0.12 | 0.62         |
| San Francisco | 2.2                        | 9.5                          | 0.40                               | 0.20  | 0.11  | 0.88         | 0.29                                | 0.13  | 0.04  | 0.96         |
| San Francisco | 1.7                        | 8.1                          | 0.34                               | -0.05 | 0.22  | 0.87         | 0.14                                | 0.16  | 0.22  | 0.92         |
| San Francisco | 2.1                        | 8.5                          | 0.52                               | 0.45  | 0.19  | 0.72         | 0.52                                | 0.29  | 0.17  | 0.72         |

also indebted to Gaylord Miller and the Data Processing Center of the University of California, Los Angeles. The remaining desk calculations were painstakingly done by Joan Bell. This work was supported in part by the Marine Research and the Scripps Tuna Oceanography research programs. Funds for the latter project were provided by the U. S. Bureau of Commercial Fisheries, Department of the Interior, Contract 008-9354.

REFERENCES

La, Fisheries Research Board, Observations of sea water temperature and salinity on the Pacific Coast of Canada, *Pacific Oceanic Group*, Vancouver, B. C., 1957-1958.  
Loomis, R. B., and J. W. Tukey, The measurement of power spectra from the point of view of communications engineering, *Bell System Tech. J.*, 37, 185-282 and 485-509, 1958.  
Loomis, N. R., On the joint estimation of the spectrum, conspectrum and quadrature spectrum of a two dimensional Gaussian process, *N. Y. Univ., Engineering Statistics Lab. Sci. Paper 10*, 1958.  
Loomis, G. W., Periodic variation in sea level induced by equatorial waves in the easterlies, *Deep-Sea Research*, 3(4), 248-252, 1956.

Groves, G. W., Day to day variation in sea level, *Meteorol. Monographs*, 2(10), 32-45, 1957.  
Haurich, R., and W. H. Munk, The pole tide, *J. Geophys. Research*, 64(12), 2373-2388, 1959.  
LaFond, E. C., Variations of sea level on the Pacific Coast of the United States, *J. Marine Research, Sears Foundation*, 2, 17-29, 1939.  
Lanning, J. H., and R. H. Battin, Random processes in automatic control, *McGraw-Hill Book Co.*, 434 pp., 1956.  
Munk, W. H., F. E. Snodgrass, and M. J. Tucker, Spectra of low frequency ocean waves, *Bull. Scripps Inst. of Oceanog. Univ. Calif.*, 7(4), 283-362, 1959.  
Munk, W. H., F. E. Snodgrass, and M. J. Tucker, Smoothing and persistence, *J. Meteorol.*, 17(1), 92-93, 1960.  
Mexico, Universidad Nacional Autonoma de Mexico Instituto de Geofisica, Observaciones mareograficas en la Republica Mexicana, Costa del Pacifico, *Boletin No. 5*, 1958.  
Nomitsu, T., and M. Okamoto, The causes of the annual variation of the mean sea level along the Japanese Coast, *Mem. Coll. Sci. Univ. Kyoto., Ser. A*, 10(3), 125-161, 1927.  
Patullo, June, W. H. Munk, R. Revelle, and E. Strong, The seasonal oscillation in sea level, *J. Marine Research, Sears Foundation*, 14(1), 88-155, 1955.



- Roden, G. I., and G. W. Groves, On the statistical prediction of ocean temperatures, *J. Geophys. Research*, 65(1), 249-264, 1960.
- Sarkisyan, A. S., Nonstationary wind driven currents in a homogeneous ocean, *Izvest. Akad. Nauk SSSR Ser. Geofiz.*, (8), 1008-1019, 1957a.
- Sarkisyan, A. S., Theory of transient wind currents in an isotropic ocean. *Izvest. Akad. Nauk. SSSR Ser. Geofiz.*, (10), 1232-1237, 1957b.
- Shoji, Daitaro, On the variation of daily mean sea levels and the Kuroshio from 1954 to 1955, *Proc. UNESCO Symposium on physical oceanography and Japan Society for the promotion of Science*, 130-136, 1957.
- Stewart, H. B., B. D. Zetler, and C. B. Taylor, Recent increases of coastal water temperature and sea level. California to Alaska, *Tech. Bull.* (3), U. S. Coast and Geodetic Survey, Washington, D. C., 1958.
- U. S. Coast and Geodetic Survey, *Surface Water Temperatures at Tide Stations, Pacific Coast, North and South America and Pacific Islands*, Fifth Ed., Washington 25, D. C., 1958.
- U. S. Coast and Geodetic Survey, Monthly sea levels at tide stations along the Pacific of the United States, unpublished.
- U. S. Geological Survey, Compilation of seawater records. Pacific Slope Basins in Oregon and the Lower Columbia River Basin (1), Washington 25, D. C., 1958.
- U. S. Weather Bureau, Monthly mean weather maps for the northern hemisphere, 1899-1958, unpublished.
- U. S. Weather Bureau, *Monthly Weather*, Washington 25, D. C., 1872-1949.
- U. S. Weather Bureau, *Climatological Data, Annual Summary*, Washington 25, D. C., 1958.
- Veronis, G., and H. Stommel, The action of variable wind stresses on a stratified ocean. *J. Marine Research*, Sears Foundation, 15, 43-75, 1958.

(Manuscript received June 2, 1960.)

# The Latest Achievements of Physical Geodesy

W. A. HEISKANEN

*Institute of Geodesy, Photogrammetry and Cartography  
Ohio State University  
Columbus, Ohio*

**Abstract.** That the accuracy of computing the undulations in the geoid is not yet as high as is desired is not due to the method itself but to the lack of gravity data. The masses of the great formations are 85 to 90 per cent isostatically compensated. The flattening ratio value,  $1/298.3$ , obtained by the satellite method has also been obtained several times by gravity and arc-measuring methods. At the Institute of Geodesy, Photogrammetry and Cartography we have been conservative; we have used only the gravimetric method and have not extrapolated gravity anomalies more than  $15^\circ$  from the computation point.

Uotila obtained by spherical harmonic analysis third-order harmonics similar to those obtained by O'Keefe from satellite data. The undulations associated with the pear-shaped form are of the order of  $\pm 15$  meters. O'Keefe's geoid heights at the north and south poles are brought about solely by the fact that the flattening value he used ( $1/299.8$ ) does not correspond to the real facts.

**Introduction.** During the last few years progress in gravimetric geodesy has been enormous. Geodesists are now interested in conducting local gravity surveys in their countries. Those companies that have made extensive local gravity surveys for exploration purposes have added their files of gravity data to the physical geodesists. The Vening Meinesz pendulum has, for several years, been in use on submarines for gravity observation at sea, and during the last two years the new method of using the sea gravimeter, invented by A. Graf in Germany, on surface vessels has been tested, particularly by J. L. Worzel's group at Columbia University. The LaCoste sea gravimeter has also been tested by other agencies with encouraging results. With the new instrumentation it has become possible to measure gravity profiles at sea, instead of at land values as has been the case in the past. The advantage of gravity profiles is obvious.

In addition, Lloyd Thompson of the Air Force Cambridge Research Center and also L. L. Nettleton's group, using a modified LaCoste gravity meter, have conducted the first tests on the feasibility of making gravity observations from aircraft. The test flights have given results which encourage us to develop this faster method for gravity survey. The invention of 'grafimeter,' as the Germans call it, and the airborne gravity

meter are now for physical geodesy of as basic significance as the invention of the Vening Meinesz pendulum was for geophysics four decades ago.

*Theses on the gravimetric method.* Before going into details I would like to mention the main theses concerning the gravimetric method.

1. The irregularities of the geoid and equipotential surfaces at higher elevations are mostly brought about by the disturbing mass anomalies, visible or invisible, shallow or deep. These irregularities cause the gravity anomalies  $\Delta g$ , which can be observed. With the aid of  $\Delta g$  we can then compute the undulations  $N$  of the geoid and the tilt, or the deflection of the vertical components  $\xi$  and  $\eta$  between the geoid and the used reference ellipsoid. Similarly, we can separate the spheropotential and geopotential surfaces in higher elevations.

2. The prerequisites of the gravimetric method are: (a) All gravity observations must be converted to the same world gravimetric system. This problem is essentially solved. (b) The whole earth's surface should be covered by a network of gravity observations.

3. The method as such is perfect. That the accuracy is not yet as high as is desired is not due to the method itself but to the lack of gravity data.

4. The earth is in almost complete equilibrium. The masses of the great formations, like high mountains and the ocean basins, are 85 to 90 per cent isostatically compensated. *Heiskanen and Vening Meinesz* [1958, p. 208, Fig. 7-5] show the seismological evidence for isostatic equilibrium and in Tables 7-1 to 7-6 (pp. 188-197) the gravimetric evidence. The light roots of the continents and the heavy antiroots of the oceans compensate the topographic masses of the continents and the oceans.

5. Of the remaining 15 to 10 per cent, the topographic masses and, of course, the irregular local or regional disturbing masses cause relatively small geoidal undulations, seldom exceeding 50 meters. Our gravity map of Ohio [*Heiskanen and Uotila*, 1956] shows that it is in regional isostatic equilibrium but that even in such a flat area there are strong local positive and negative anomalies.

6. If no isostatic equilibrium prevailed, the undulation values  $N$ , as Helmert computed 80 years ago, would exceed 400 meters under the big continents. It is clear, then, that isostatic equilibrium is the best 'benefactor' of the geodesists. Without its help the separation of the reference ellipsoid and the geoid would be many times more difficult and would require much more detailed gravity observations than are necessary on an earth in equilibrium.

7. The gravimetric method is certainly the most powerful branch of geodesy because it can relate the different continents to the same world geodetic system; it is the only method capable of giving the detailed shape of the geoid not only on the continents but also over the oceanic areas. It is also the only method which can yield the gravity field and its irregularities, as well as the shape of the spheropotential surfaces with and without the effect of the centrifugal force to any elevations, knowledge of which is of vital significance particularly in satellite geodesy.

8. The gravimetric method, or physical geodesy, requires accurate knowledge only of the equatorial radius  $a$  and the equatorial gravity  $\gamma_e$ . The flattening value  $\alpha$  of the meridian is, however, of minor significance. If we change the flattening value, the gravity formula, and with it the gravity anomalies, will also change. Consequently, we get different sets of gravity anomalies

TABLE 1. Some Flattening Values of the Meridian Obtained by Different Methods

| $\alpha$ | Year | Author        | Method               |
|----------|------|---------------|----------------------|
| 1:298.1  | 1856 | Clarke        | Arc-measuring        |
| 1:298.3  | 1901 | Helmert       | Gravity              |
| 1:298.3  | 1929 | Heiskanen     | Arc-measuring        |
| 1:298.3  | 1938 | Heiskanen     | Gravity              |
| 1:298.3  | 1942 | Krassowsky    | Arc-measuring        |
| 1:298.4  | 1912 | Bowie         | Gravity              |
| 1:298.4  | 1942 | Krassowsky    | Arc-measuring        |
| 1:298.38 | 1959 | O'Keefe       | Artificial satellite |
| 1:298.32 | 1959 | Eckels        | Artificial satellite |
| 1:298.24 | 1959 | Cook          | Artificial satellite |
| 1:298.4  | 1959 | King-Hele     | Artificial satellite |
| 1:298.29 | 1959 | Jacchia       | Artificial satellite |
| 1:298.3  | 1959 | Adopted value | Artificial satellite |

and different sets of  $N$  values for every flattening value  $\alpha$ . The geocentric radius of a point on the geoid would be the same in both cases, however, provided that the earth's surface were reasonably well surveyed gravimetrically.

Table 1 shows some flattening values obtained by different methods. This table indicates that the flattening value 1:298.3 obtained by the artificial satellite method has been previously obtained several times by gravity and measuring methods.

The situation is exactly similar to that in computing topographic elevations. If we use sea level as a reference surface, we get, for instance, different values for the elevation of a point in Ohio than we would get if we used the level of Lake Erie as a reference. However, the radius of any point in Ohio from the earth's center would be the same in both cases.

9. It is important to remember that the gravimetric quantities  $N$ ,  $\xi$ , and  $\eta$ , are almost absolute and are essentially independent of the equatorial radius and of the deflections of the vertical at the initial point of the geodetic datum.

*The Columbus studies.* The significance of the improved gravity anomaly field was well understood in Columbus. In 1952 we made a detailed suggestion of the transoceanic gravity profile that would be required to provide the needed gravity information. This map was sent to the appropriate agencies for review. For the southern hemisphere we suggested profiles at an average distance of 10° of latitude and for



ern hemisphere at an average distance of this program would have taken little more than 2 years to complete. Because due attention was not paid to this proposal we are now much farther along than we were 8 years ago insofar as filling the large gaps in our knowledge of the southern oceans is concerned. On the other hand, the fast methods of obtaining gravity at sea from surface ships will very soon give gravity data from all oceans. At the needed amount of gravity data we have spent much time and effort in contacting geodetic institutions, universities, independent geodesists, and oil companies operating in different parts of the world. Also, analyzing the needed material has taken quite a bit of time. The fact that we now get data from 55 different countries, either directly or indirectly, has enormously helped not only our efforts in Columbus but also the efforts of several government agencies who have used this material for similar computations. The maps showing the gravity anomalies available in Columbus are shown in *Kaula and Hirvonen* [1958].

*Different approaches.* The goal of all scientists in physical geodesy who have to determine the dimensions of the earth, to convert the existing geodetic systems to one world system, and to determine the gravity anomaly field and the dimensions of the equipotential surfaces at any elevations, of course, the same. As to the methods, there are different opinions, as is also natural. I mention some examples: Bomford and his group prefer to use only the classic astrogeodetic method; the satellite men seem to be inclined to overestimate the geodetic significance of the information obtained from the period of the satellite revolution and the precession of the orbit and perigee of the satellite orbits. Some scientists try to combine the results of the astrogeodetic method, satellite geodesy, and the gravimetric method [*Kaula*, 1959a]. Other scientists are inclined to fill even continent-wide geographically unsurveyed areas with gravity anomalies estimated by the methods of statistical mathematics. Others trust, perhaps too much, the possibilities of the development of  $\Delta g$  and  $N$  in spherical harmonics.

The Columbus group has deliberately used only the gravimetric method. We have been conservative. We have not ventured to extrapo-

late gravity anomalies into unsurveyed areas farther than  $15^\circ$  from the computation point. With this method only 10 per cent of the observed gravity is used to represent a square. Basic studies in this respect were made by *Hirvonen* [1957] in Columbus in 1953. *Kaula* [1957], of the Army Map Service, has continued these studies.

For still larger areas the isostatic anomaly zero is, in our opinion, the most representative. We have applied the Stokes integral and/or the Vening Meinesz formula to the observed gravity anomalies and interpolated and extrapolated these to reasonable distances. It is true we have also developed (U. A. Uotila, unpublished report) the gravity anomalies in spherical harmonics, Legendrians, sectorials, and tesserals of the gravity anomalies to the eighth order. The figures are also ready for the sixteenth order. We are however, skeptical as to whether the terms of higher degree and order have any physical meaning or whether they are rather purely mathematical computation figures without reliable physical meaning. When more gravity observations are available from the southern hemisphere, where the largest gaps exist, the significance of the development of the gravity anomalies to the spherical harmonics will indeed increase. All of us know that this method is very convenient if sufficient gravity data are available.

*Different reduction methods.* Opinions also differ concerning the best reduction method to be used. Several new suggestions have been made. The terms 'model earth' of de Graaff Hunter, 'quasi geoid' of Molodensky, and the 'telluroid' of Hirvonen, as well as several investigations carried out by other scientists, show that we are obviously not very far from the advent of a new reduction method. Until this happens, however, it is only reasonable to use the best of the existing methods.

The essential difference between the classic and the modern way of thinking is this. In the classic method, which is also used in all practical computations in Columbus, at least so far, we reduce the obtained gravity to sea level, or geoid, and compare this value with the theoretical value of gravity at the geoid. In the new methods the theoretical gravity will be 'uplifted' to the observational point, and the comparison between the observed and theoretical gravity will be



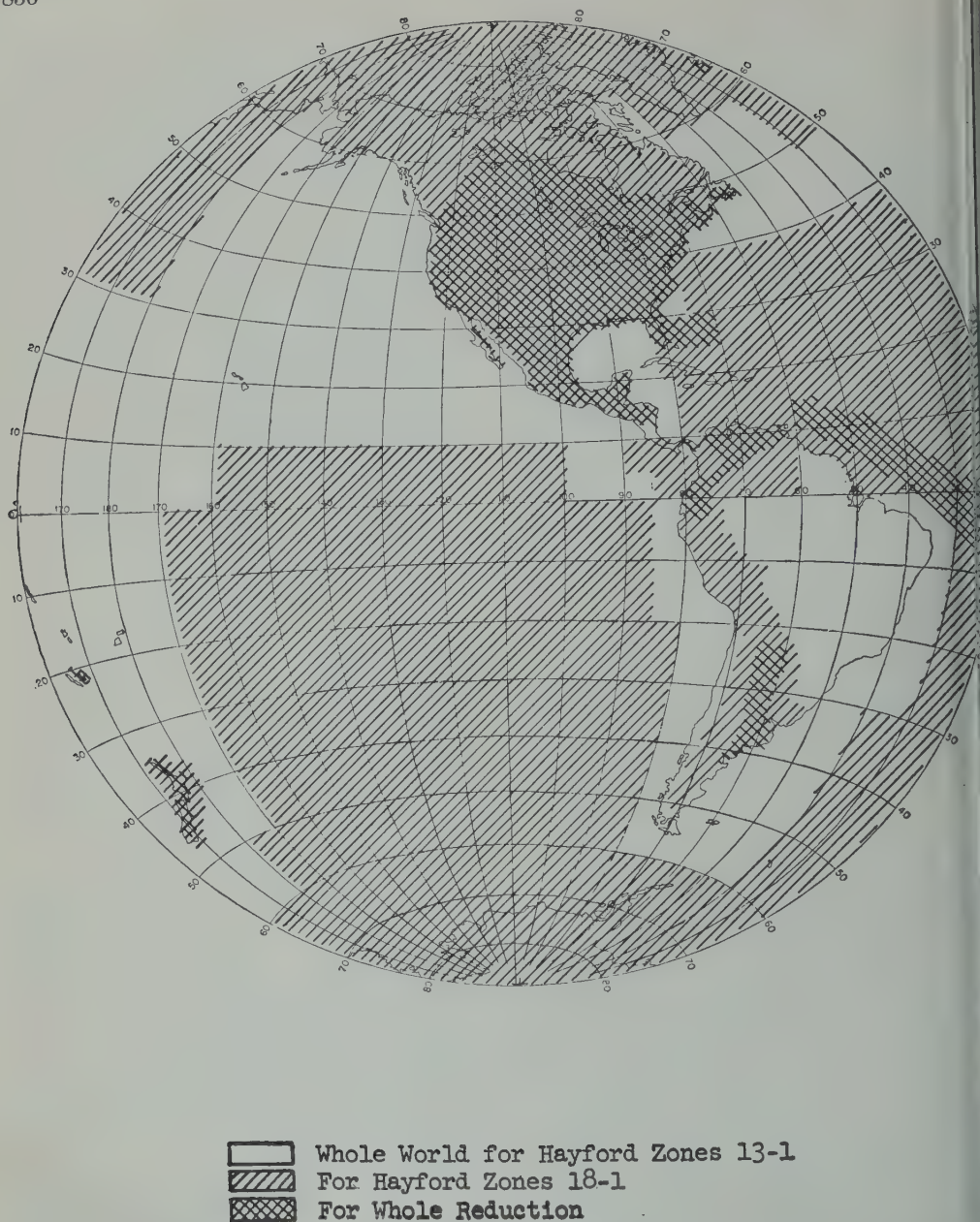





Fig. 1. Isotatic reduction maps for  $T = 30$  km available December 31, 1959 (western hemisphere).

made at the observation point. Then we will get the gravity anomalies at the observation point itself. I suggested four years ago to Dr. Hirvonen, my colleague since 1930, that he study the new methods. He has indeed succeeded well. He published (May), an investigation, 'New Theory of

the Gravimetric Geodesy' [Hirvonen, 1960] produced earlier as two technical reports of the Institute of Geodesy, Photogrammetry and Topography of Ohio State University. According to him we do not now need the geoid; we can use the telluroid, a three-dimensional reference



-  Whole World for Hayford Zones 13-1  
 For Hayford Zones 18-1  
 For Whole Reduction

Isostatic reduction maps for  $T = 30$  km available December 31, 1959 (eastern hemisphere).

According to several geodesists he goes too far in this respect. Since 70 per cent of the earth's surface is covered by oceans and since the geoid and the telluroid are the same surface over the oceans, the geoid is therefore the most representative. In addition, about 15 per cent of

the earth's surface is flat land, lower than 500 meters, where the difference  $N_1$  between the geo- and spheropotential surfaces and the undulations  $N$  of the geoid are essentially the same. There remains, however, 15 per cent of the areas of high mountains and rough topography near



sea level that are adjacent to the coast lines, where, in practice, a larger difference exists between the classic and modern methods of treating the data.

Whatever method we use we have some difficulties in the rough mountain areas. If we use the free-air anomalies, as has been suggested for modern reduction, we have to find some reasonable way of getting rid of the disturbing effect of the very irregular free-air anomalies in the neighborhood of the computation point.

At the same time that we in Columbus have made basic studies on the new method, we have spent a great deal of time in carrying out the isostatic reductions, because we intend to use the isostatic gravity anomalies in the next computation of the geoid. In fact, for a large part of the world we now have isostatic mean anomalies for  $1^\circ \times 1^\circ$  squares; the normal thickness of the earth's crust, 30 km, is used in the computations. In addition, about 10,000 individual gravity values are reduced isostatically. Figures 1 and 2 show the areas which have been either completely or partly isostatically reduced.

The isostatic reduction has several advantages: First, it gives more representative gravity anomalies than any other method; second, the isostatic anomaly zero is a good representative for the unsurveyed areas; third, the isostatic anomalies are of vital significance in the geophysical studies of the earth. To make the isostatic reduction as easy as possible, we have prepared and drafted world maps from which either the whole reduction or the reduction of different Hayford zone groups can be obtained. When we add the isostatic reduction to the usual Bouguer correction we get the isostatic anomalies. It is true that in areas of rugged topography a terrain correction has to be considered; fortunately, for about 85 per cent of the earth's surface no terrain correction is needed.

In every reduction of the gravity anomalies in which we think the mass changes, the equipotential surface changes as well. Therefore, the Bowie reduction, or indirect effect  $\Delta g$  must be added to the isostatic anomalies to get the correct anomalies on the computed geoid, or co-geoid. In this way we get correct gravity anomalies and, using the Stokes formula, the right shape of the co-geoid. Then it is a matter of only one multiplication to transfer the co-geoid to the actual geoid.

Some geodesists have claimed that the method described above and used by us and by several other scientists is wrong in principle, but they have apparently misunderstood the computation of isostatic reduction. When the indirect effect is considered, the method used by us gives, in principle, as accurate a result as any other method.

*Comparison between the different geoids.* Three years ago the first gravimetrically computed geoid in the western hemisphere, the 'Columbus geoid,' was presented [Heiskanen, 1959]. The Army Map Service has since computed other geoids using a different extrapolation method, somewhat different material, and different flattening values  $\alpha$ . The computed geoids differ from one another, of course, because we have used the flattening value of the international ellipsoid,  $\alpha = 1:297.0$ , and the Army Map Service [Kaula, 1959a] has used the value  $\alpha = 1:298.49$ , or  $\alpha = 1:298.3$ . It is relatively simple to convert the  $N$  values from one flattening value to another. When we did this in Columbus it turned out that the Army Map Service geoid and the Columbus geoid agree rather well in most areas of the northern hemisphere where there is a rather good network of gravity stations existing. For the southern hemisphere we have not yet compared the Columbus geoid with the Army Map Service geoid. In the Columbus geoid, made any computations of  $N$ , because in these areas no gravity observations, or very few, exist.

We hope before the Helsinki Assembly to be able to compute the 'second edition' of the Columbus geoid, using the isostatic gravity anomalies. Then it will be obviously possible to extend the geoid to some areas of the southern hemisphere. Theoretically, there should not be spherical harmonics  $P_{1,0}$ ,  $P_{1,1} \cos \lambda$ ,  $P_{1,1} \sin \lambda$ ,  $P_{2,1} \cos \lambda$ , and  $P_{2,1} \sin \lambda$ . When, in fact, these spherical harmonics are relatively small but not zero, they are obviously brought about mostly because the available gravity data are inadequate. Therefore, Hirvonen [1960] and Uusikallio [1960] have balanced the gravity field, making these spherical harmonics zero. The corrections have been, of course, added to the unsurveyed and poorly surveyed areas but not to areas where an adequate gravity anomaly net exists. The largest corrections are, of course, in the Pacific region, where in some cases they are as high as +9 mgal; in the western Atlantic corrections are +6 mgal, in Africa +3 mgal,

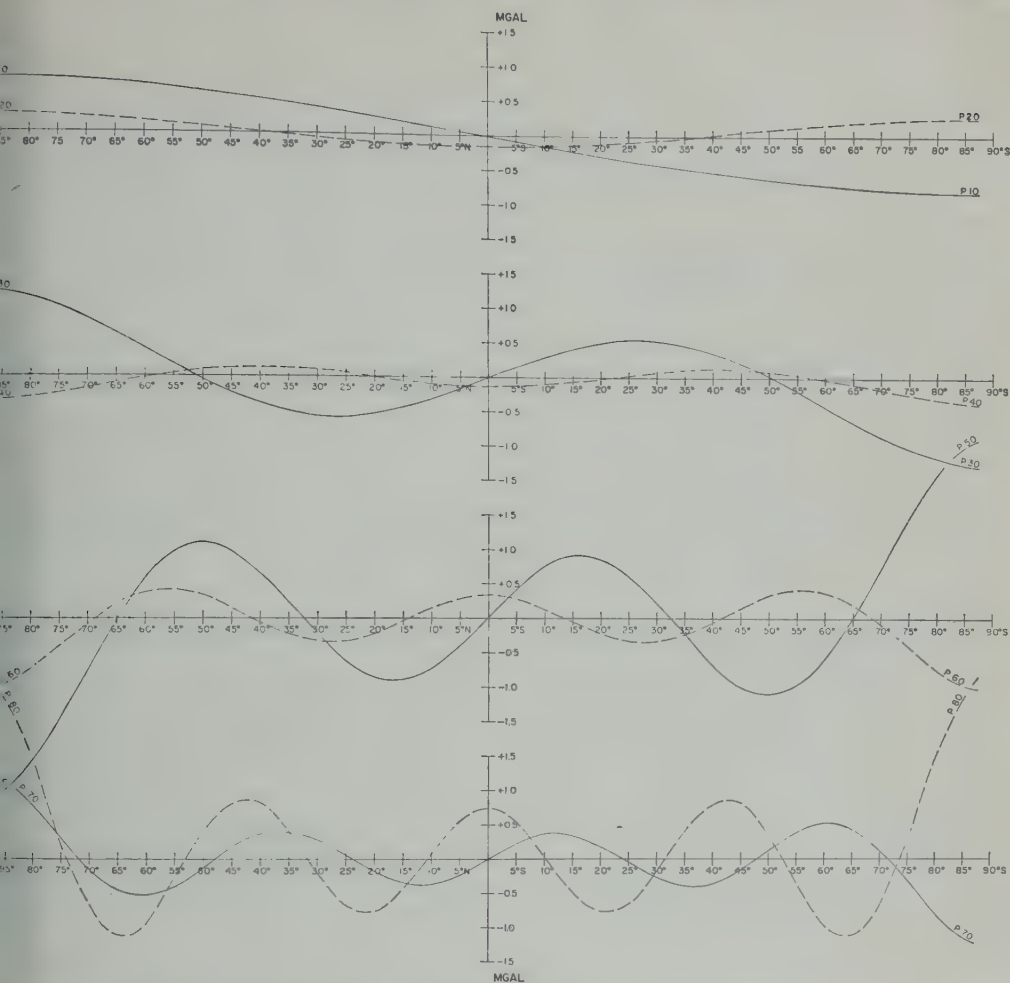


Fig. 3. Gravity anomalies developed in spherical harmonics (Legendrians) for flattening  $\alpha = 1:297.0$ . Computed by U. Uotila, July 1959, Institute of Geodesy, Photogrammetry and Cartography, Ohio State University.

Indian Ocean  $-3$  mgal, etc. The effect of the corrections on the computed gravimetric section of the vertical is, in general, very small; at no computed point do they exceed  $\pm 1$  mgal.

In the spring of 1959 Uotila computed the first-order Legendrian harmonics for the corrected gravity anomalies, using the flattening values  $1:297.0$  and  $1:298.3$ . The results are to be seen in Figures 3, 4, 5. The third-order harmonics are shown in Figures 3 and 4.

By the artificial satellite method O'Keefe [1959, Fig. 2, p. 2391] has obtained third-order harmonics similar to those Uotila (unpublished

report) obtained from gravity observations. Both these facts indicate that the earth is perhaps a trifle pear-shaped, which is to be expected. In the northern hemisphere there are the most continents, in the southern hemisphere the most oceans. Since the isostatic equilibrium is short about 10 to 15 per cent, both hemispheres can be different. However, the undulations of the geoid brought about by the pear-shaped form is only of the order of  $\pm 15$  meters—rather small but not negligible—when compared with the systematic 40-meter undulations of relatively large areas [Heiskanen, 1957].

O'Keefe's publication on the pear-shaped



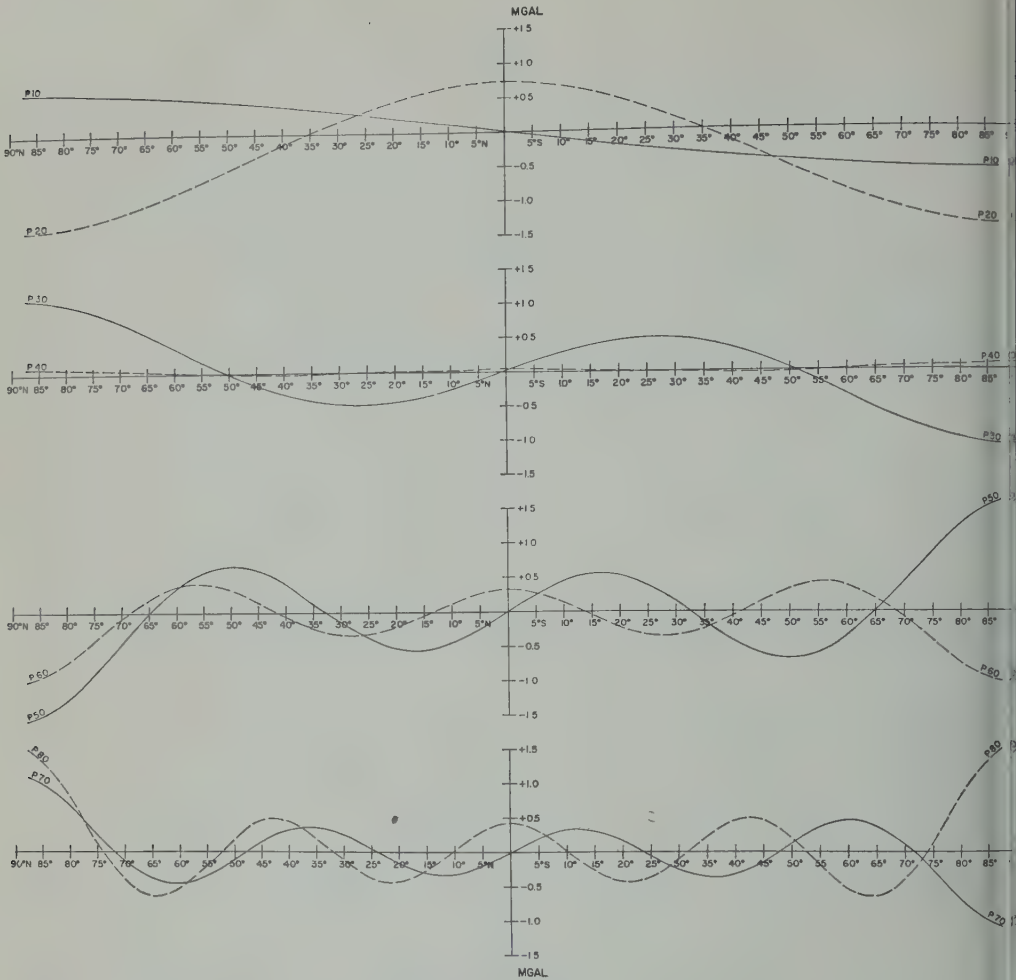


Fig. 4. Gravity anomalies developed in spherical harmonics (Legendrians) for flattening  $\alpha = 1:292$ . Computed by U. Uotila, July 1959, Institute of Geodesy, Photogrammetry and Cartography, Ohio State University.

earth has also found critics. *Brenner, Fulton, and Sherman* [1960] say:

By means of a special numerical integration carried out for nine hundred revolutions, it is shown that the oblateness of Earth causes long period changes in the shape, size and inclination of the osculating orbit of the Vanguard I satellite (1958 Beta two). Unless many determinations of these parameters are made per revolution, the effect of an Earth with equatorial symmetry can imitate in period, amplitude and phase the effect which might be expected from a "pear-shaped" potential. This result, which has been verified by independent analysis using an appropriate perturbation theory,

shows that published evidence does not support the hypothesis that Earth's mass distribution is asymmetric with respect to its equatorial plane.

*O'Keefe* [1959] claims that 'the basic hypothesis of geodesy stated by Vening Meinesz. Heiskanen calls for an extremely smooth gravity field for the earth as a whole, apart from local irregularities. From satellite measurements the zonal harmonics of orders 2, 3, and 4 it is shown that the actual roughness is about an order of magnitude greater than that demanded by the basic hypothesis of geodesy.' This basic hypothesis is the assumption that the elliptical

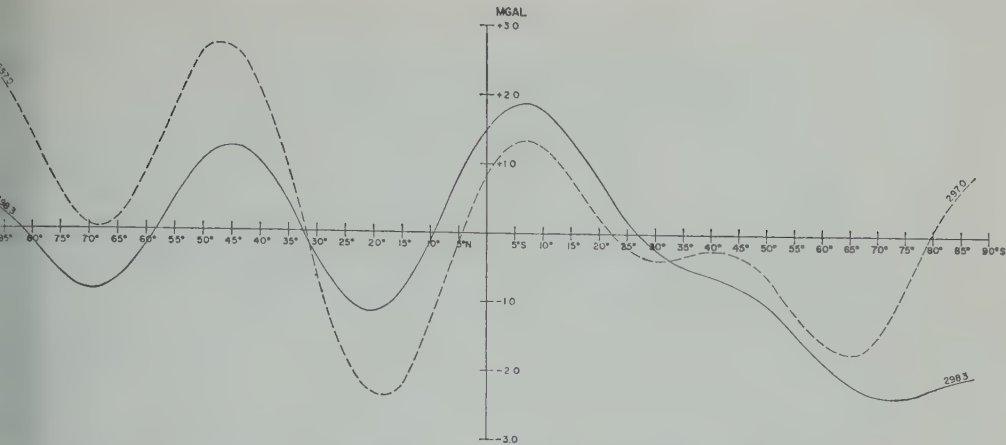


Fig. 5. Sum of the eight first Legendrians of the gravity anomalies for the flattening values of 297.0 298.3. Computed by U. Uotila, July 1959, Institute of Geodesy, Photogrammetry and Cartography, State University.

evolution should represent the equilibrium of the rotating earth.

In fact, we have not claimed that complete hydrostatic equilibrium prevails. On the contrary, *Heiskanen and Vening Meinesz* [1958] say: 'We can be sure, however, that the basic hydrostatics has outlived its usefulness and that it is no longer a sound basis for geodesy.' O'Keefe has assumed that the flattening value 1:299.8 corresponds to the hydrostatic equilibrium, and he uses this flattening value when he computes

the spherical harmonics of the gravity anomalies and the undulations  $N$  of the geoid. Because the hydrostatic equilibrium is *not* complete, one cannot use this value, and Figure 2, page 2931, of *O'Keefe's* [1959] publication will not explain anything. It indicates only that the huge geoid heights of  $-73$  meters at the South Pole and  $+38$  meters at latitude  $40^{\circ}\text{N}$  are brought about solely by the fact that the flattening value used by him does not correspond to the real facts.

Since the difference between the equatorial

TABLE 2. Potentialities of Physical Geodesy and Satellite Geodesy

|   | Physical Geodesy             | Satellite Geodesy                  |
|---|------------------------------|------------------------------------|
| Equatorial radius $a$   | Can help                     | Can                                |
| Flattening $f$  | Can                          | Can                                |
| Spherical harmonics of $\Delta g$   |                              |                                    |
| Legendrians   | Any degree                   | 2nd, 3rd, 4th, 6th degree          |
| Sectorials  | Any order                    | Cannot                             |
| Tesserals   | Any order                    | Cannot                             |
| Undulations $N$   | At any points, also at sea   | Only low harmonics                 |
| Deflections $\xi, \eta$   | At any points, also at sea   | Nobody will try to use this method |
| Shape of the sphero- and geopotential surfaces  | At any elevations and points | Cannot                             |
| Gravity field $\gamma_h, g_h, \Delta g_h$   | At any elevations            | Cannot                             |
| Gravitation field $\Gamma_h, g_h, \Delta g_h$   | At any elevations            | Cannot                             |
| Connect the continents  | Can                          | Can                                |
| Structure of the earth's crust  | Can                          | Cannot                             |
| Only prerequisites: Sufficient gravity anomaly field at the earth's surface and all anomalies in the same World Gravimetric System. |                              |                                    |

and polar radius brought about by the flattening is 21,500 meters, the change of the flattening value corresponding to 1 unit of the denominator of the flattening corresponds to 70 meters. If O'Keefe had used, for instance, the flattening value 1:301, he would have got almost twice as big 'undulations' of the geoid.

I accept, however, one sentence of Dr. O'Keefe: 'The difficulty of applying the gravimetric method is simply a difficulty rising from the lack of measurements of wide areas of the earth.' We have, however, drawn from this fact different conclusions. Vening Meinesz has devoted a large part of his time to getting oceanic gravity surveys going, and I myself have tried for one decade to get different agencies interested in gravity observations in the unsurveyed areas. If we had used for the gravity surveys one small part of the money which has been used for the tracking of artificial satellites, the gaps in the gravity anomaly field would, some years ago, have been filled and we would now know the detailed shape of the geoid very well.

We are all certainly very glad that different methods exist for the determination of the dimensions and shape of the earth. Table 2 gives an idea of the problems which physical geodesy and satellite geodesy can solve.

*Acknowledgments.* I wish to thank the Directorate of the Air Force Cambridge Research Center for financial support of these studies and the Aeronautical Chart and Information Center and the Army Ballistic Missile Agency for their endorsement of our studies. It is also a pleasure to express my thanks to Hans Baussus, Bela Szabo, James Peoples, and Owen Williams for their continued interest, help, and cooperation.

## REFERENCES

- Brenner, J. L., R. Fulton, and N. Sherman, Symmetry of the earth's figure, *ARS Journal*, March, 278-279, 1960.
- Heiskanen, W. A., The Columbus geoid, *Trans. Am. Geophys. Union*, 38, 841-848, 1957.
- Heiskanen, W. A., On the gravimetric undulations of the geoid and the deflections of the vertical, *Geofis. pura e appl.*, 41, 1-18, 1958.
- Heiskanen, W. A., and F. A. Vening Meinesz, *The Earth and Its Gravity Field*, McGraw-Hill Book Co., New York, 1958.
- Heiskanen, W. A., and U. A. Uotila, Gravity survey of the State of Ohio, *Publ. No. 6, Inst. Geod., Photogram. Cartog., Ohio State Univ.*, 1956.
- Hirvonen, R. A., The precision obtainable by the eclipse method, *Publ. of Inst. Geod., Photogram. Cartog., Ohio State Univ.*, 3 pp, 1957.
- Hirvonen, R. A., New theory of the gravimetric geodesy, *Publ. Isostatic Inst. of Intern. Assoc. Geodesy*, No. 32, 50 pp., 1960.
- Kaula, W. M., Accuracy of gravimetrically computed deflections of the vertical, *Trans. Am. Geophys. Union*, 38, 297-305, 1957.
- Kaula, W. M., Reconciliation of Stokes' function and astro-geodetic geoid determinations, *J. Geophys. Research*, 64, 61-71, 1959a.
- Kaula, W. M., Statistical and harmonic analysis of gravity, *J. Geophys. Research*, 64, 2401-2422, 1959b.
- O'Keefe, J. A., Zonal harmonics of the earth's gravitational field and the basic hypothesis of geodesy, *J. Geophys. Research*, 64, 2389-2399, 1959.
- Uotila, U. A., Determination of the shape of the geoid, *Publ. Inst. Geod., Photogram. Cartog. No. 7, Ohio State Univ.*, pp. 90-97, 1957.
- Uotila, U. A., Investigations on the gravity field and shape of the earth, *Inst. Geod., Photogram. Cartog., No. 10, Ohio State Univ.*, 72 pp., 1960.

(Manuscript received May 23, 1960.)

# Hiran Geodesy and Photographic Observations

ALEXANDER CORPACIUS

*Department of Astronomy  
Georgetown University, Washington 7, D. C.*

**Abstract.** New formulas are developed for computing the azimuth and distance between two Hiran ground stations with data obtained from the photographs; orientation calibration of the cameras is not necessary. Thereby, a method of determining the orientation of Hiran network is obtained.

This paper is a logical development of a pre-study [Corpaci, 1960] of flare triangulation with photographic observations.

Assume two Hiran ground stations  $A$  and  $B$ , provided with a ballistic camera. The cameras are oriented toward an airplane  $S$  which is making a standard line-crossing. Simultaneously with the Hiran measurement of successive distances  $r = SA$  and  $r' = SB$ , lights on the airplane are flashed and are photographed by both cameras. From Hiran measurements of  $r$ ,  $r'$  and coordinates of the flare and of three reference stars, highly accurate azimuth and distance between ground stations can be computed.

According to the study mentioned above, the following equations,

$$x_0 = x_1 D_1 + x_2 D_2 + x_3 D_3 \quad (1)$$

$$y_0 = y_1 D_1 + y_2 D_2 + y_3 D_3 \quad (2)$$

$$1 = D_1 + D_2 + D_3 \quad (3)$$

to the values of the so-called dependences  $D_1, D_2, D_3$ , where

$x_0, y_0$  = plate coordinates of a flare.

$x_1, y_1, x_2, y_2, x_3, y_3$  = plate coordinates of three reference stars.

On the other hand, the formulas by Comrie [1955],

$$\begin{aligned} & \bar{\alpha} + \sum D_i (\alpha_i - \bar{\alpha}) \\ & - \tan \bar{\delta} \left[ \sum D_i (\alpha_i - \bar{\alpha}) (\delta_i - \bar{\delta}) \right] \\ & - (\alpha_0 - \bar{\alpha}) (\delta_0 - \bar{\delta}) \end{aligned} \quad (4)$$

$$\begin{aligned} \delta_0 &= \bar{\delta} + \sum D_i (\delta_i - \bar{\delta}) \\ &+ \frac{\sin 2\bar{\delta}}{4} \left[ \sum D_i (\alpha_i - \bar{\alpha})^2 - (\alpha_0 - \bar{\alpha})^2 \right] \\ &(i = 1, 2, 3) \end{aligned} \quad (5)$$

might be used for computing the right ascension  $\alpha_0$  and declination  $\delta_0$  of the same flash by means of  $\alpha_i, \delta_i$  of the same reference stars and dependences  $D_1, D_2, D_3$ . ( $\bar{\alpha}, \bar{\delta}$  are the approximative right ascension and declination of the center of the plate, which can be read directly from a sky chart.)

Now, a Cartesian space system  $OXYZ$ , located in the center of a given spheroid is adopted. When axis  $OX$  is parallel to the line of equinox,  $OZ$  is parallel to the line of poles, and  $OY$  is perpendicular to the plane  $XOZ$  and directed toward the east, then the direction cosines,

$$\begin{aligned} u &= \cos \alpha_0 \cos \delta_0; & u' &= \cos \alpha'_0 \cos \delta'_0 \\ v &= \sin \alpha_0 \cos \delta_0; & v' &= \sin \alpha'_0 \cos \delta'_0 \\ w &= \sin \delta_0; & w' &= \sin \delta'_0 \end{aligned} \quad (6)$$

of the rays joining the camera stations  $A$  and  $B$  with the flare  $S$ , can be computed by means of (4) and (5). Moreover, camera locations  $A(X_A, Y_A, Z_A)$  and  $B(X_B, Y_B, Z_B)$  and flare positions  $S(X, Y, Z)$  are related by formulas

$$\begin{aligned} X &= X_A + ru; & X &= X_B + r'u' \\ Y &= Y_A + rv; & Y &= Y_B + r'v' \\ Z &= Z_A + rw; & Z &= Z_B + r'w' \end{aligned} \quad (7)$$



After  $X$ ,  $Y$ ,  $Z$  are eliminated, the following conditional equations are obtained:

$$\Phi = ru - r'u' - X_B + X_A = 0$$

$$\psi = rv - r'v' - Y_B + Y_A = 0 \quad (8)$$

$$\chi = rw - r'w' - Z_B + Z_A = 0$$

When we assume that the camera station  $A$  is located in a geodetic point which belongs to a given datum, the only unknowns of the conditional equations (8) are  $X_B$ ,  $Y_B$ ,  $Z_B$ ;  $u$ ,  $v$ ,  $w$  and  $u'$ ,  $v'$ ,  $w'$  are functions of the directly measured plate coordinates of the flash and of the reference stars.

With (6), the conditional equations (8) can be written

$$\Phi = r \cos \alpha_0 \cos \delta_0 - r' \cos \alpha'_0 \cos \delta'_0 - X_B + X_A = 0$$

$$\psi = r \sin \alpha_0 \cos \delta_0 - r' \sin \alpha'_0 \cos \delta'_0 - Y_B + X_A = 0 \quad (9)$$

$$\chi = r \sin \delta_0 - r' \sin \delta'_0 - Z_B + Z_A = 0$$

After linearizing the conditional equations (9) and replacing  $\alpha_0$ ,  $\delta_0$  and  $\alpha'_0$ ,  $\delta'_0$  by directly measured plate coordinates according to Comrie's formulas, we get the following new conditional equations derived by Hiran technique and photographic observations:

$$\begin{aligned} a_0 v_{x_0} + b_0 v_{y_0} + a_1 v_{x_1} + b_1 v_{y_1} + a_2 v_{x_2} \\ + b_2 v_{y_2} + a_3 v_{x_3} + b_3 v_{y_3} + a'_0 v'_{x_0} + b'_0 v'_{y_0} \\ + a'_1 v'_{x_1} + b'_1 v'_{y_1} + a'_2 v'_{x_2} + b'_2 v'_{y_2} \\ + a'_3 v'_{x_3} + b'_3 v'_{y_3} - dX_B + \phi_0 = 0 \end{aligned} \quad (10)$$

$$\begin{aligned} e_0 v_{x_0} + f_0 v_{y_0} + e_1 v_{x_1} + f_1 v_{y_1} + e_2 v_{x_2} + f_2 v_{y_2} \\ + e_3 v_{x_3} + f_3 v_{y_3} + e'_0 v'_{x_0} + f'_0 v'_{y_0} \\ + e'_1 v'_{x_1} + f'_1 v'_{y_1} + e'_2 v'_{x_2} + f'_2 v'_{y_2} \\ + e'_3 v'_{x_3} + f'_3 v'_{y_3} - dY_B + \psi_0 = 0 \end{aligned} \quad (11)$$

$$\begin{aligned} g_0 v_{x_0} + h_0 v_{y_0} + g_1 v_{x_1} + h_1 v_{y_1} + g_2 v_{x_2} + h_2 v_{y_2} \\ + g_3 v_{x_3} + h_3 v_{y_3} + g'_0 v'_{x_0} + h'_0 v'_{y_0} \\ + g'_1 v'_{x_1} + h'_1 v'_{y_1} + g'_2 v'_{x_2} + h'_2 v'_{y_2} \\ + g'_3 v'_{x_3} + h'_3 v'_{y_3} - dZ_B + \chi_0 = 0 \end{aligned} \quad (12)$$

where the coefficients, which can be easily computed, and the unknowns have the following meaning:

$\left. \begin{matrix} a, b, e, f, g, h \\ a', b', e', f', g', h' \end{matrix} \right\} = \begin{matrix} \text{coefficients dependent} \\ \text{on the measured plate coor} \\ \text{dinates.} \end{matrix}$

$\Phi_0, \psi_0, \chi_0 = \text{free (absolute) terms.}$

$\left. \begin{matrix} v_{x_0}, v_{y_0} \\ v_{x'_0}, v_{y'_0} \end{matrix} \right\} = \begin{matrix} \text{unknown increments} \\ \text{directly measured} \\ \text{coordinates } x_0, y_0, x'_0, y'_0 \\ \text{of flare, corresponding to} \\ \text{camera stations } A \text{ and } B, \text{ re} \\ \text{spectively.} \end{matrix}$

$\left. \begin{matrix} v_{x_i}, v_{y_i} \\ v_{x'_i}, v_{y'_i} \end{matrix} \right\} (i = 1, 2, 3) = \begin{matrix} \text{unknown increments} \\ \text{directly measured} \\ \text{coordinates } x_i, y_i, x'_i, y'_i \\ \text{of three reference} \\ \text{corresponding to cam} \\ \text{stations } A \text{ and } B, \text{ re} \\ \text{spectively.} \end{matrix}$

$dX_B, dY_B, dZ_B = \begin{matrix} \text{unknown corrections} \\ \text{assumed space coordi} \\ \text{ates } X_B, Y_B, Z_B \text{ of cam} \\ \text{station } B. \end{matrix}$

Summarizing the results obtained so far we see that:

1. With one flash, the unknowns  $dX_B$ ,  $dY_B$ ,  $dZ_B$  may be computed directly from equations 9; with more than one flash, a least squares treatment of the data is possible (10), (11), and (12) as conditional equations.
2. The accuracy of conditional equations 11, and 12 is reflected by Comrie's equations 4 and 5, which are valid up to the second approximation in  $\Delta\alpha_0$  and  $\Delta\delta_0$ .
3. After this computation of camera location, distance and azimuth between Hiran station  $A$  and  $B$  may be derived according to standard procedures.

#### REFERENCES

- Comrie, L. J., Note on the reduction of photographic plates, *J. Brit. Astron. Assoc.*, Nov. 1929.
- Corpaci, A., Rigorous adjustment of flare triangulation with photographic observations, Presented at the Forty-First Annual Meeting of the American Geophysical Union, Section of Geodesy, Washington, D. C., April 27-30, 1960. Submitted for publication in *Bull. géodésique*, Paris.

(Manuscript received April 11, 1960; revised May 17, 1960.)

# Horizontal Movement in the Earth's Crust

C. A. WHITTEN

*U. S. Coast and Geodetic Survey  
Washington 25, D. C.*

**Abstract.** The conventional method for determining horizontal movement in the earth's crust has been to reobserve networks of triangulation and compare the coordinates of the adjusted results. A new method of analysis of reobservations is presented. The changes in the angles in a network indicate the presence of strain or deformation within the crust. This type of analysis will also indicate small displacements which may occur along fault lines in an area of seismic activity. Results of the application of this technique to resurveys along the San Andreas Fault in California are given in graphical form.

The problem of measuring small horizontal movements on the surface of the earth in areas of seismic activity has generally been solved by observing a triangulation network. If there have been changes, the difference in the observations will indicate the magnitude of the movements. Because of the complexity and interlocking characteristics of a triangulation network, the usual practice for calculating these shifts of position has been to compute geographic coordinates of the points for each set of observations and then to compare the coordinates. There is always the chance that limitations imposed by inevitable errors of observation or weaknesses in the network itself may produce accumulated shifts that exceed any conceivable ground movement. Also a stable base line must always be available as a reference for the network. Any small shift in one of the reference points can produce fictitious shifts of position of the other points (references.)

Geodesists have been aware of these features as a basic problem but have endeavored to overcome these weaknesses by establishing a regular schedule of reobservation over carefully monumented points. Studies based on triangulation data have shown the general pattern of movement along the San Andreas fault zone. In an effort to obtain the maximum amount of information from the observations, a different type of analysis is presented. These new techniques are not intended to replace the classical methods but rather to supplement them and to provide a more definitive interpretation.

In areas adjacent to the San Andreas Fault where major forces within the earth are producing these horizontal movements, we may assume that small rectangular areas near the fault line will be gradually deformed into parallelograms. The axes of these geometrical areas must be considered to be parallel to the fault. Precise triangulation observations can be used to compute the amount and rate of this deformation. The quantity can be expressed in terms of seconds of arc referred to directions perpendicular to the fault line.

Frequently there has been evidence of slippage or minute displacements along the fault. If some of the lines in a triangulation net cross the fault, an unknown for this displacement can be introduced into the analysis.

The mathematical development is based on the differential concepts. In Figure 1, lines from  $P$  to  $A$ ,  $B$ , and  $C$  represent the directions to be measured. Subscripts 1 and 2 indicate the relative changes between the points when the area is subjected to the previously mentioned forces. The central point,  $P$ , may be moving in an absolute sense, but the differential effects are adequate to determine the local deformation. It is assumed that the direction of the movement is parallel to the fault. This is consistent with the history of the movements along the fault.

In Figure 2,  $\delta$  represents the angle of deformation measured in the direction perpendicular to the fault line.  $\delta$  is small and may be treated as a differential.  $Q_1$  and  $Q_2$  are imaginary points on a line through  $A$  parallel to the fault line, and

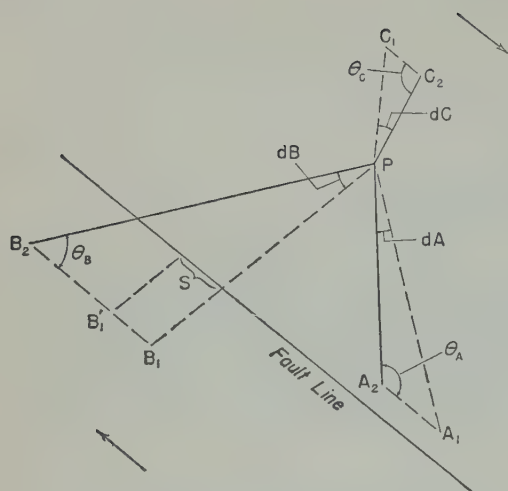


Fig. 1. Deformation in directions and relative displacement of points.

$Q_1$  is on a line through  $P$  perpendicular to the fault line.

$$A_1 A_2 = Q_1 Q_2 \quad (1)$$

$$\sin \delta = \frac{Q_1 Q_2}{P Q_2} = \frac{A_1 A_2}{P Q_2} \quad (2)$$

$$\frac{\sin dA}{A_1 A_2} = \frac{\sin \theta_A}{P A_1} \quad (3)$$

$$\sin \theta_A = \frac{P Q_1}{P A_2} \quad (4)$$

From (2) and (3)

$$\sin dA = \frac{P Q_2 \cdot \sin \delta \cdot \sin \theta_A}{P A_1}$$

$P Q_1$  and  $P Q_2$  are essentially equal, as are  $P A_1$  and  $P A_2$ . Thus, for the small angles involved  $\sin dA = \sin \delta \cdot \sin^2 \theta_A$  or  $dA = \delta \sin^2 \theta_A$ . Similarly  $dC = \delta \sin^2 \theta_C$ .

In figure 1, the line,  $P B$ , crosses the fault line.  $dB = \delta \sin^2 \theta_B +$  increment due to  $S$ .  $S$ , the slippage along the fault, is uniform in magnitude rather than proportional as  $A_1 A_2$ ,  $B_1 B_2$ ,  $C_1 C_2$ , etc. Therefore the angular change at  $P$  in the direction to  $B$  due to  $S$  is dependent upon the length,  $P B$ , and  $\theta_B$ .

The sine of the increment due to "S" is  $S \cdot \sin \theta_B / PB$ , or the increment expressed in seconds due to "S" is  $S \cdot \sin \theta_B / PB \cdot \sin 1''$ .

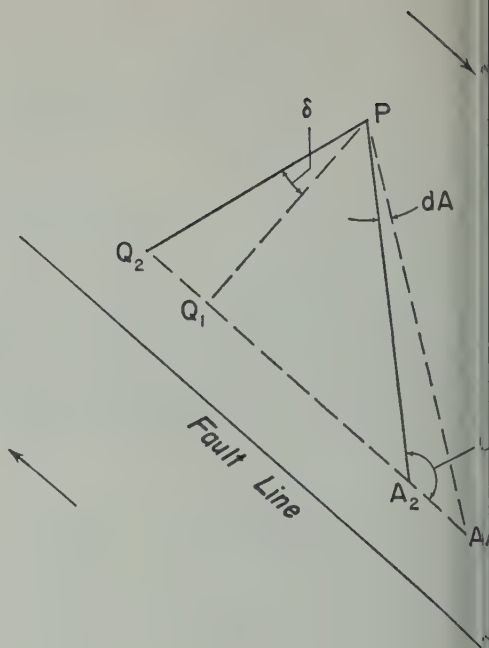


Fig. 2. Deformation angle  $\delta$ .

Therefore

$$dB = \delta \sin^2 \theta_B + S \cdot \frac{\sin \theta_B}{PB \sin 1''}$$

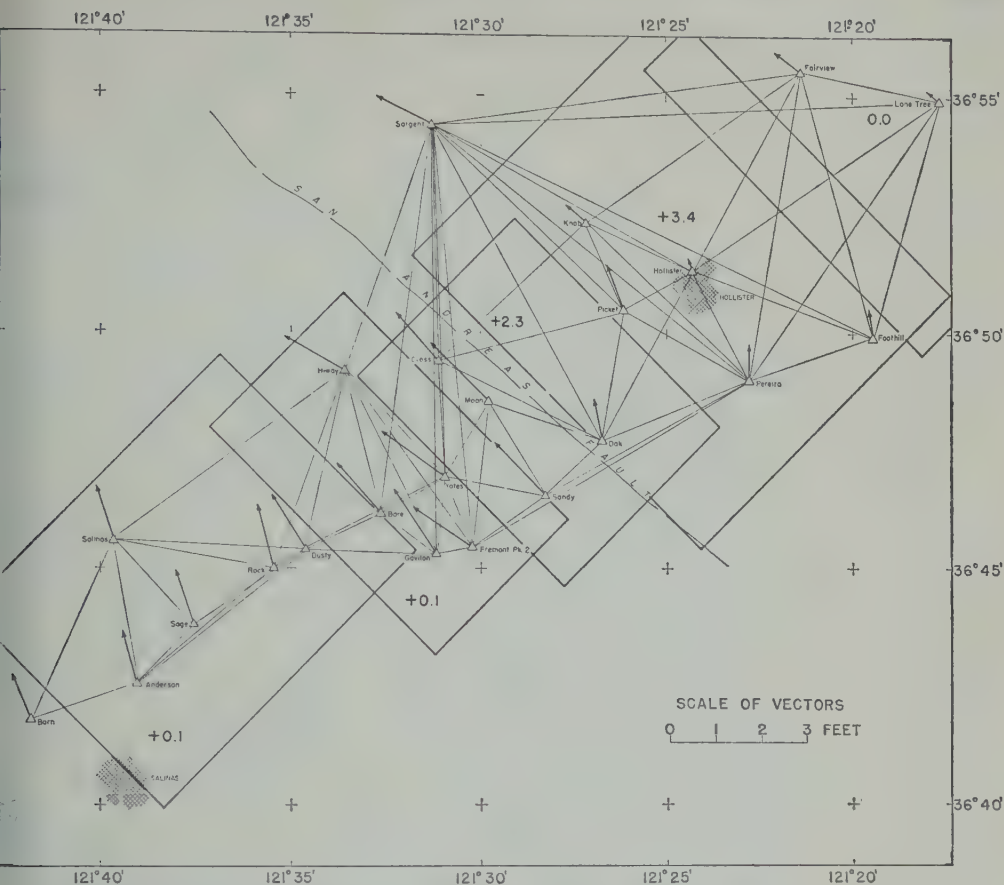
In the above equation for the angular change of the line,  $P B$ , the second term provides means for computing the slippage, represented by  $S$ . The first term of the equation, it will be noted, is the same as in the general form where all points are on the same side of the fault line.

These equations cannot be applied directly to observational data. When two lists of direct observations made at different times are compared, another unknown,  $K$ , must be introduced. This is a station constant, merely a matter of rotation of one set of observations without any change to angles. Thus, the general observation equation takes the form

$$V = K + \delta \sin^2 \theta + \frac{S \cdot \sin \theta}{\sin 1''}$$

The third term is used only when a line of observation crosses the fault line.

Theoretically, the unknowns can be computed from the observations at any one point. However,



3. Network in vicinity of Hollister. Vectors indicate movement. Rate of deformation in seconds of arc per 10-year period is shown by figures in blocks. Slippage per 10 years is 10 cm.

the small quantities to be determined are of the same order of magnitude as observational errors. By grouping four or five points together and solving for a common value of  $\delta$  (and  $S$ , if available), considerable statistical advantage is gained. The results are essentially the same as those that could be obtained by computing weighted means of the values as determined from the single-point observations.

This new technique has been applied to several sets of observations in California. The results are presented in summary form with the rate of deformation expressed as a 10-year average. The slippage or displacement quantities are weighted means over the same 10-year period. The diagrams of the networks, with movement expressed as vectors, are ready for direct comparison.

The basic material shown in Figures 3, 4, and 5 has been published [Whitten, 1956]. The vectors shown in Figure 3 represent the horizontal movement which took place between surveys of 1930 and 1951. The reference or fixed points for this survey are 10 to 15 miles northeast of the limit of the diagram. The vectors shown in Figure 4 represent horizontal movement between surveys of 1932 and 1951. The reference or fixed points are also 10 to 15 miles northeast of the limit of the diagram. The vectors in Figure 5 indicate horizontal movement between surveys of 1941 and 1954. The reference or fixed points are those on the northeastern side of the net where no vectors are shown.

The trend is the same in the three figures. The vectors on the southwest side of the fault are larger than those on the northeast side. The



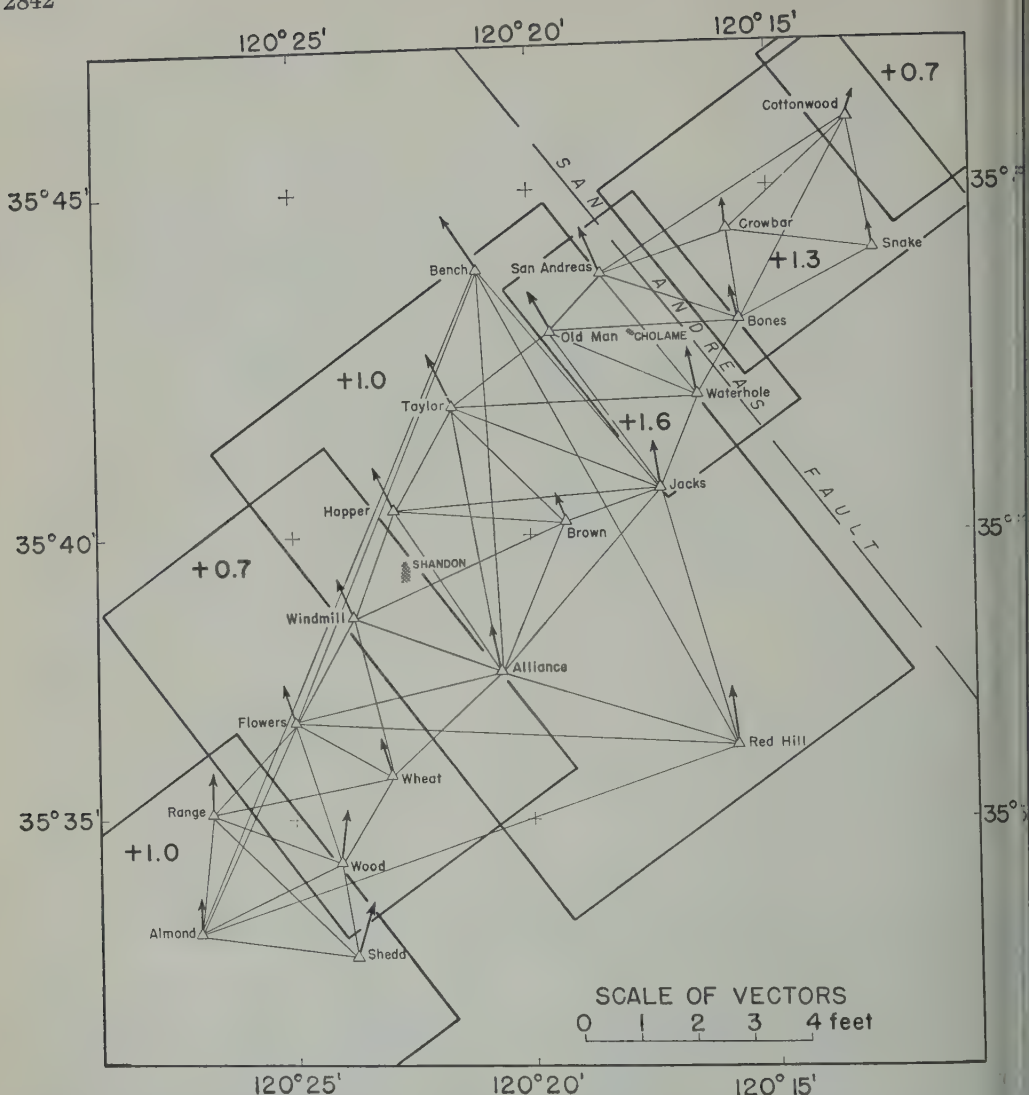


Fig. 4. Network in vicinity of Cholame. Vectors indicate movement. Rate of deformation in sec of arc per 10-year period is shown by figures in blocks. Slippage per 10 years is 3 cm.

relative size of the vectors between the two sides indicates the differential movement. A sharp difference in the vectors on opposite sides of the fault indicates displacement or slippage.

It is significant that the pattern and the rates of deformation are in fair agreement even though the surveys are widely separated, the distance between the two most widely separated being more than 400 miles. We hope to be able to apply this type of analysis to the large

amount of data which has accumulated over past 30 years and is in the archives of the U. S. Geological and Geodetic Survey, but this will require considerable effort, time, and man power.

In some localities along the fault there is positive evidence of more or less continuing slippage. In such places, small survey patterns can be established and used as a frame for measuring the displacement. One such place is Hollister. Two very small survey figures,

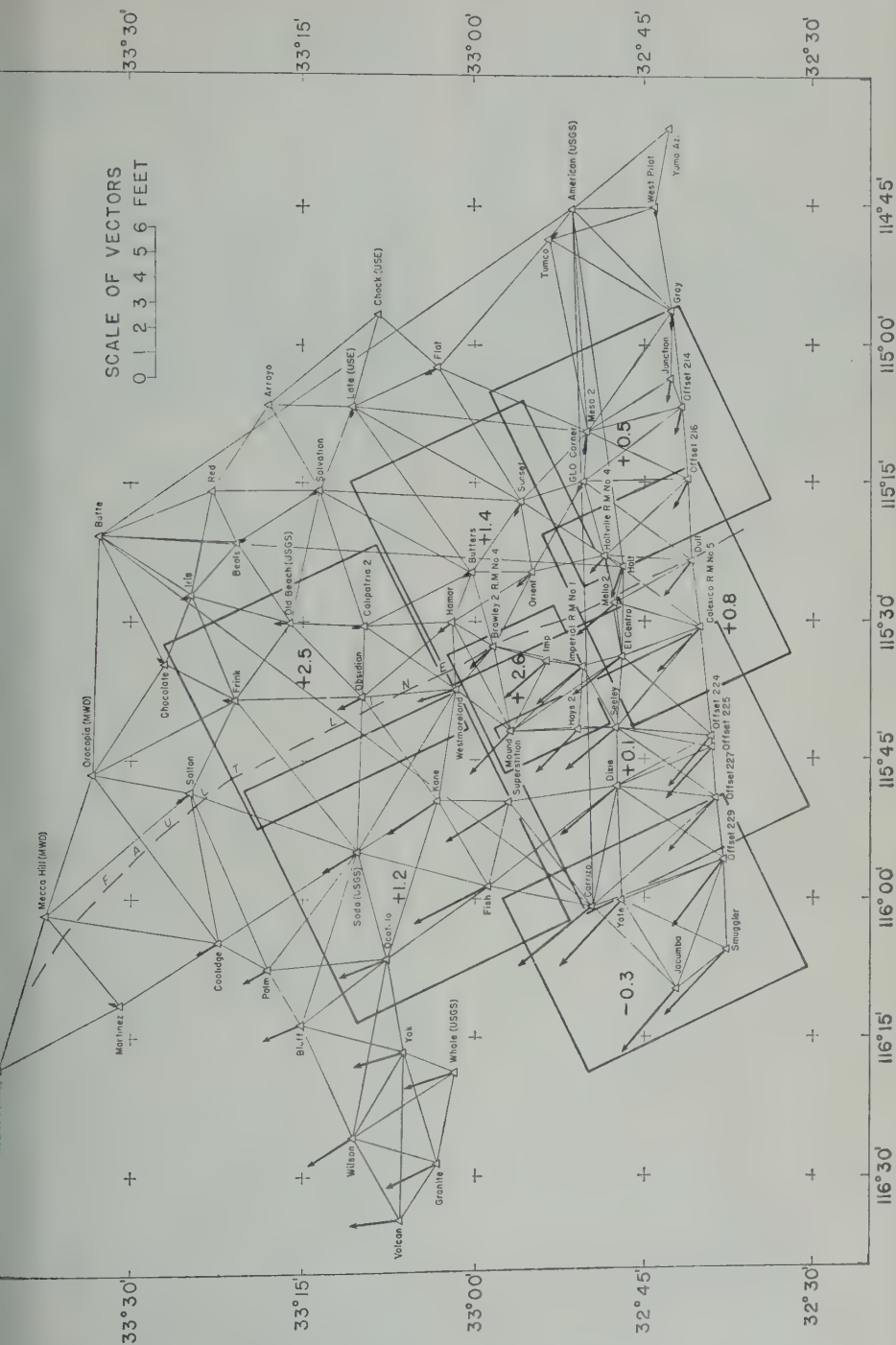


Fig. 5. Network in vicinity of El Centro. Vectors indicate movement. Rate of deformation in seconds of arc per 10-year period is shown by figures in blocks. Slippage per 10 years is 4 cm.

only a few hundred meters in extent, were placed straddling the fault. The first observations were made in April 1957. Repeat observations made in September 1958 disclosed a slippage of  $20 \pm$  mm. Seismologists and engineers have set up means of making direct measurements at more closely spaced intervals of time. The fault line is so well defined that scales can be used almost like verniers, with direct readings to fractions of millimeters. This is a very special case, but it indicates what can be done. The cost of making measurements of this type over closely spaced points is a fraction of that required for repeating the more extensive networks. However, one must keep in mind the fact that only displacement is measured in these small patterns. If studies of deformation over broad areas are desired, the larger surveys must be made to provide the observational data for such calculations.

## REFERENCES

- Harada, Y., Re-survey of the southwestern part of Japan after the great Nankaidō earthquake of 1946, *Bull. Geograph. Survey Inst. (Japan)* 1, 1952.
- Kasahara, Keichi, The nature of seismic origin inferred from seismological and geodetic observations, 1 *Bull. Earthquake Research Tokyo Univ.*, 35 (3), 1957.
- Whitten, C. A., Horizontal earth movement, vicinity of San Francisco, California, *Trans. Am. Geophys. Union*, 29, 318-323, 1948.
- Whitten, C. A., Horizontal earth movement in California, *The Journal, Coast and Geodetic Survey*, No. 2, April 1949.
- Whitten, C. A., Measurements of earth movements in California, *Calif. Dept. Nat. Resources Div. Mines, Bull. 171*, 1955.
- Whitten, C. A., Crustal movement in California and Nevada, *Trans. Am. Geophys. Union*, 393-398, 1956.

(Manuscript received May 31, 1960.)

# Simultaneous Adjustment of Angular and Distance Measurements

FREDERICK F. CEELY, JR.

*U. S. Coast and Geodetic Survey  
Washington 25, D. C.*

**Abstract.** The development and use of electronic distance-measuring instruments has posed a problem that many of our triangulation networks have become a mixture of triangulation, trilateration, and traverse, especially in the areas where control is needed for the interstate highway program. A method of adjusting triangulation, trilateration, and traverse simultaneously is presented in this paper. The formulas, methods, and techniques were devised in an attempt to make an adjustment that gives a satisfactory solution with minimum distortion to both angular and distance measurements. A presentation of several approaches to the problem is made with emphasis given to the application of these methods to an intermediate-sized electronic computer.

The problem of adjusting angular and distance measurements simultaneously has existed many years. Specifically, the use of modern electronic distance-measuring equipment in some recent surveys has produced a mixture of triangulation, trilateration, and traverse. An example of the type of problem now being encountered is shown in Figure 1.

In adjusting triangulation, the classical method of adjustment of coordinates on the ellipsoid is used. The three conventional unknowns are rotation, shift of latitude, and shift of longitude. The particular method requires many calculations and we are using an IBM 650 magnetic-electronic computer.

Information needed as 'input' includes the observed positions of the fixed control points, observed directions, and approximate estimated positions of new points to be adjusted (these positions are adequate for this).

The mathematician in charge of the adjustment, after inspecting a sketch of the survey, assigns identification code numbers to all the points. These numbers are then placed in the directions adjacent to the conventional station names. Our programs are written to solve the equations in a numerical sequence, so that the three unknowns at station 1 are solved by the three unknowns at stations 2 and 3. Therefore, the manner in which the stations in the scheme are numbered will determine the order of elimination. The mathematician

must exercise careful judgment in the assignment of these number codes to keep the number of multiplications or 'drag' to a minimum. Preliminary inverse position computations are made over the lines indicated by the lists of directions. Then the absolute term of each ob-

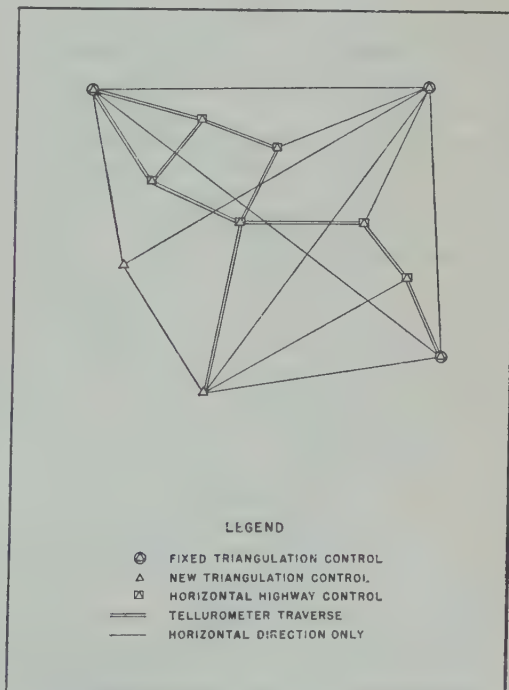


Fig. 1. Example of interstate highway survey.



servation equation is computed by comparing the inversed azimuth of the line with the azimuth of the line computed through the list of directions. From these inverses and the proper absolute terms observation equations for each line can be computed.

The direction observation equations have the general form [Whitten, 1957]

$$V_{ik} = Z_i - \frac{2062.65}{S} \left( \frac{\sin \alpha_{ik}}{B_i} \delta \phi_i + \frac{\cos \alpha_{ki}}{H_k} \delta \lambda_i + \frac{\sin \alpha_{ki}}{B_k} \delta \phi_k - \frac{\cos \alpha_{ki}}{H_k} \delta \lambda_k \right) + N$$

where

$i$  is the point from which the direction is observed.

$k$  is the point to which the direction is observed.

$V_{ik}$  is the correction at  $i$  to the direction  $k$ .

$Z_i$  is the rotation unknown for the arbitrary initial of a list of directions (making the sum of the residuals at a given point equal to zero).

$\alpha$  is the azimuth of the line between  $i$  and  $k$ .

$S$  is the length of the line between  $i$  and  $k$ .

$\delta \phi_i$  and  $\delta \phi_k$  are the unknown shifts in latitude of  $i$  and  $k$ .

$\delta \lambda_i$  and  $\delta \lambda_k$  are the unknown shifts in longitude of  $i$  and  $k$ .

$N$  is the difference between the assumed azimuth and the azimuth from the list of directions from  $i$  and  $k$ .

$$B = \frac{(1 - e^2 \sin^2 \phi)^{3/2}}{a(1 - e^2) \arc 1''}, \quad H = \frac{(1 - e^2 \sin^2 \phi)}{a \cos \phi \arc 1''}$$

The condition equation for azimuth has the general form [Whitten, 1957]

$$0 = -\frac{2062.65}{S} \left( \frac{\sin \alpha_{ik}}{B_i} \delta \phi_i - \frac{\cos \alpha_{ik}}{H_i} \delta \lambda_i + \frac{\sin \alpha_{ki}}{B_k} \delta \phi_k - \frac{\cos \alpha_{ki}}{H_k} \delta \lambda_k \right) + N$$

where

$N$  = (inversed azimuth) - (fixed or Laplace azimuth)

The condition equation for length has the general

form

$$0 = \frac{1}{B_i} \cos \alpha_{ik} \delta \phi_i + \frac{1}{H_k} \sin \alpha_{ki} \delta \lambda_i + \frac{1}{B_k} \cos \alpha_{ki} \delta \phi_k - \frac{1}{H_k} \sin \alpha_{ki} \delta \lambda_k$$

where

$N$  = (inversed length) - (fixed or measured length)

The above formulas are the standard ones that have been used for several years in the adjustment of triangulation where we are concerned only with corrections to directions.

The first attempt to incorporate an equation that would give corrections to lengths was a modification of a method published by [1946]. Belote introduced a system of adjusting a traverse net by computing condition equations for length, azimuth, latitude, and longitude which included terms for a scale correction to be applied to the measured distances.

We theorized that the tellurometer was used for most of our measurements had a constant error as well as a scale error. A modification of the general conditions for length was made by adding two terms; one, a constant, and the second, representing the length correction,  $lS/10,000$ . This modified condition equation has the general form

$$0 = \frac{1}{B_i} \cos \alpha_{ik} \delta \phi_i + \frac{1}{H_k} \sin \alpha_{ki} \delta \lambda_i + \frac{1}{B_k} \cos \alpha_{ki} \delta \phi_k - \frac{1}{H_k} \sin \alpha_{ki} \delta \lambda_k + l + \frac{lS}{10,000} + N$$

where  $l$  is a constant term to be applied to all lengths in a section.  $lS/10,000$  is a scale correction to be applied to all lines in a section.

This type of length condition was introduced into the normal terms, and several adjustments were made using this method. Although the results were satisfactory, no two solutions gave the same value for the constant term; and the lengths in a traverse section were changed by the same scale correction—either all increased or all decreased.

l decreasing. In addition to these undesirable features, we ran into difficulty on several occasions when we tried to solve the normal equations.

The Cholesky method for the solution of normal equations is used. In this method the square root of each leading term is divided into the remaining terms of the reduced normals for each equation, giving terms which are the geometric means of the two terms with which we are dealing. This method is very familiar in the standard Gauss-Dolittle method. When we encountered a traverse in which the lines ran nearly east and west or nearly north and south, the nature of the observation equations, being functions of the azimuth of the line, would give coefficients very close to zero. Upon normalizing these equations by dividing by the sine of the azimuth, trying to solve them, we would develop lead terms that were close to zero. When we tried to divide these zeros as divisors, there were overflows and machine stops. This was a clear indication that the normal equations were unstable and a new approach was necessary.

In the next attempt an observation equation for length was introduced. This was derived by multiplying the standard condition equation for length and dividing all the terms by  $S \sin 1''$  setting it equal to a  $V$  as shown in the following:

$$\frac{2062.65}{S} \left( \frac{1}{B_i} \cos \alpha_{ik} \delta \phi_k \right. \\ \left. + \frac{1}{H_k} \sin \alpha_{ki} \delta \lambda_i + \frac{1}{B_k} \cos \alpha_{ki} \delta \phi_k \right. \\ \left. - \frac{1}{H_k} \sin \alpha_{ki} \delta \lambda_k \right) + \frac{206265N}{S}$$

Now, having an observation equation for length measurements and one for distance measurements, we are interested in the mechanics of the system. Earlier it was stated that angles and absolute terms were computed for each line in the solution. When these inverses were made, all the necessary functions needed for the computation of each term in the observation equations were in the 'output.' Similar computations are made over lines involving measured angles to form length observation equations. Separate programs are needed to form the co-

efficients for the two types of observation equations. After the cross multiplications have been made for each set of equations they are added together to form one set of normal equations. These are solved by the method described earlier. In addition to these normal equations, many projects need additional conditions for length and azimuth, which are inserted into the normals.

A conventional back solution is made, producing the unknowns of the solution. These values are substituted into the original observation equations for the computation of the  $V$ 's or residuals. Although residuals from the direction observation equations represent seconds of correction to the list of directions, the residuals from the length equations must be divided into 206265, the reciprocal of the sine of  $1''$ , resulting in a proportional change to the lengths. For example,  $1''$  is approximately 1 part in 200,000 and  $10''$  is about 1 part in 20,000. The unknowns representing shifts in latitude and longitude are applied to the approximate geographic positions to obtain the final or adjusted positions which are used to make final inverses over all the lines in the solution.

This type of adjustment has been used primarily for second-order highway control surveys. Since the horizontal observations are considered to be more accurate than the tellurometer lengths, a system of weights was applied. Assuming a standard error of 1.5 seconds on the angles and the standard error of one part in 40,000 in the tellurometer lengths, and applying a ratio of weights based upon the inversed square of the standard errors, a weight ratio of 10:1 is being used for the majority of our adjustments.

*Conclusion.* The adjustment of surveys which include triangulation, traverse, and trilateration by the method of direction-observation equations and length-observation equations began in the summer of 1958 and has continued to this day. In that time we have adjusted over 3500 miles of traverse, averaging about 2.8 miles per leg. The average corrections to the lengths vary from 1 part in 26,000 to 1 part in 120,000, with an over-all average correction of 1 part in 48,000. The corrections to the angles in most cases met the standards for second-order specifications.

*Acknowledgment.* The writer is one of a group of geodesists who worked on this project under the leadership of Mr. Charles A. Whitten, who contributed most of the ideas expressed in this paper.

## REFERENCES

Belote, E. S., Adjustment of traverse, *Trans. Am. Geophys. Union*, 27, 307-315, 1946.

Gale, L. A., Simultaneous adjustment of triangulation and trilateration: An investigation of the effect of scale errors on the determination of the meridian arc length, *J. Geophys. Research*, 65, 472-477, 1960.

Rainsford, H. F., *Survey Adjustments and the Method of Least Squares*, Frederick Ungar, New York, 1957.

Whitten, C. A., Distortion in area nets of triangulation, *Festschr. C. F. Baeschlin*, Zurich, 1959.

(Manuscript received May 31, 1960.)

# Sub-Bottom Reflection Measurements on the Continental Shelf, Bermuda Banks, West Indies Arc, and in the West Atlantic Basins<sup>1</sup>

JOHN EWING, BERNARD LUSKIN, ARCHIE ROBERTS, AND JULIUS HIRSHMAN

*Lamont Geological Observatory  
Columbia University  
Palisades, New York*

**Abstract.** During a recent cruise of *R.V. Vema*, a short survey was made with the sub-bottom depth recorder in the vicinity of the Hudson Canyon between the 40-fathom and the 100-fathom contours. The results show that on this part of the continental shelf the upper layers are approximately horizontal, whereas the deeper ones dip seaward at a slightly greater angle than the bottom. Several reflecting horizons were observed at depths as great as 400 feet below the bottom, and many were traced for distances of the order of 10 to 20 miles. Where the canyon cuts through the sediments of the shelf, several of the sub-bottom horizons outcrop on the canyon walls, sometimes correlating with benches or knees. Some of the horizons at the outcrops are practically horizontal, in particular those on the northeast wall; others, notably the deep ones on the southwest wall, are sloping, tending to follow the contour of the canyon wall. Sediment cores were taken in a few locations, providing some information about the reflecting horizons (reflectors).

On Challenger Bank and on Plantagenet Bank near Bermuda, no sub-bottom reflectors were found. On the West Indies Arc near Puerto Rico and the Virgin Islands, shallow reflectors were seen in some areas.

During a run from New York to Bermuda to Puerto Rico, sub-bottom reflectors were seen on the echo-sounder fathogram at least 40 per cent of the time. The best results were obtained on the topographic rises, particularly on the outer rise north of the Puerto Rico Trench.

## INTRODUCTION

The sub-bottom depth recording equipment used in this survey was developed at the Lamont Geological Observatory after a design produced at the Woods Hole Oceanographic Institution [Knott and Hersey, 1956]. The equipment and some of the results obtained with it in Long Island Sound, in Chesapeake Bay, and in the Beagle Channel have been described previously [Eckmann, Roberts, and Luskin, 1959]. Basically, the equipment operates much like a standard echo sounder except that the acoustic source is an electric spark or a gas explosion instead of the usual crystal or magnetostrictive transducer. Both the spark and the explosion produce a signal with good low-frequency content. This provides a sounder capable of penetrating several hundred feet of water and sediment and allows sub-bottom horizons to be mapped in the same manner that the ocean floor is mapped by conventional echo sounders. In

the work described here, the spark source was used, operating with a repetition rate of 2 sec<sup>-1</sup> and utilizing time-variable gain in the recording. This produces two duplicate records, one at low

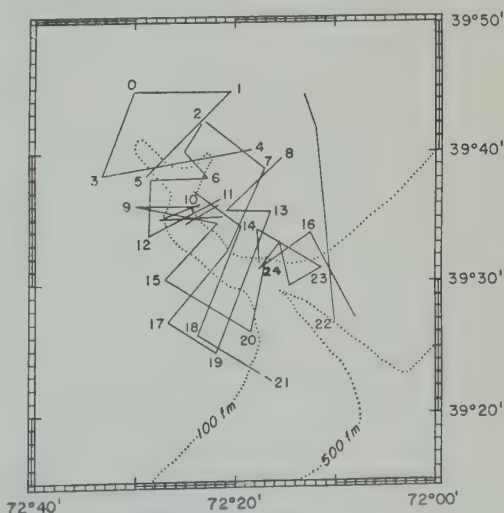


Fig. 1. SDR traverses of the Hudson Canyon.



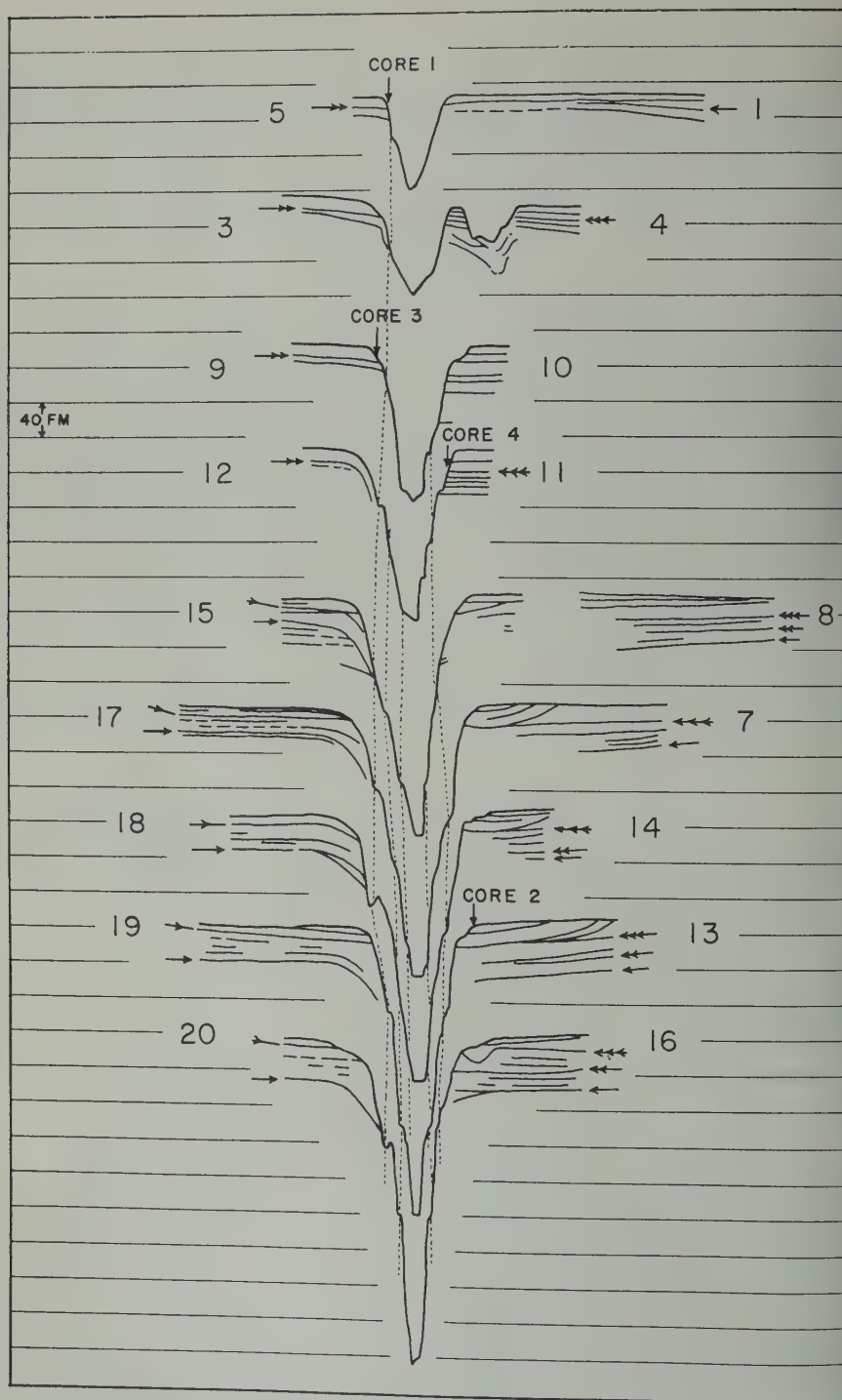


Fig. 2. Sections across the Hudson Canyon reconstructed from SDR records. Identifying numbers e.g. 3-4, are drawn on the grid line corresponding to 80 fm for each profile.

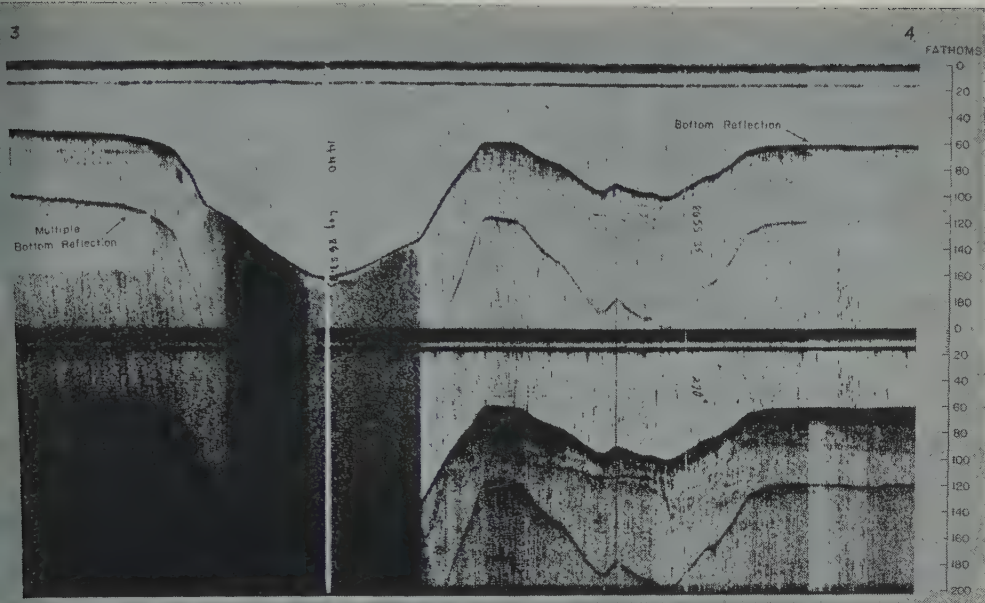


Fig. 3. SDR record, profile 3-4.

ne at high sensitivity. In several of the , only the low-sensitivity record is shown. spark electrode unit was towed 40 to 50 behind the ship, and the receiving hydro- 60 to 90 feet behind the spark. The aver- eed of the ship during the survey was 5 This produces a record in which the verti- aggeration is about 10:1. Sample records own in Figures 3 to 7.

he deep basin areas the records were made precision depth recorder (PDR) [Luskin, , Ewing, and Landisman, 1954; Luskin Roberts, 1955] which utilizes an AN/ 1 Sonar sounding set operating at a fre- y of 12 kc/s.

#### PRESENTATION AND DISCUSSION OF RESULTS

Continental Shelf-Hudson Canyon. The sur- rea and tracks along which recordings made are shown in Figure 1. For con- ce in identifying particular profiles, per- course changes on the track chart are red 0 to 24, and the corresponding num- re used to label the profiles. The 100- and hom contours are shown by dotted lines. re 2 shows nine profiles of the canyon ed so that from top to bottom each suc-

ceeding crossing was made farther down the canyon. The grid lines are 40 fathoms apart (based on sound velocity of 4800 ft/sec). No depth corrections have been made, either for water velocity or sediment velocity. Nearby seismic measurements [Ewing, Worzel, Steen- land, and Press, 1950; Oliver and Drake, 1951] give velocities in the upper part of the sedimen- tary column of 5500 to 5800 ft/sec. Hence, the actual thicknesses of the various layers of sedi- ments are probably 15 to 20 per cent greater than shown in the sections.

The few horizons that can be correlated with reasonable certainty from one profile to another are indicated by arrows of similar form. Un- fortunately, it is not possible to say whether the horizons on opposite sides of the canyon can be correlated, because the track around the north end of the canyon produced a record of poor quality. Additional study will be made in the near future in an attempt to provide better ties between the two sides of the canyon and to test the extent of continuity of the re- flectors over wide areas of the shelf. Some pos- sible correlations of structural benches and knees have been indicated by connecting them with dashed lines. Also indicated in the diagram

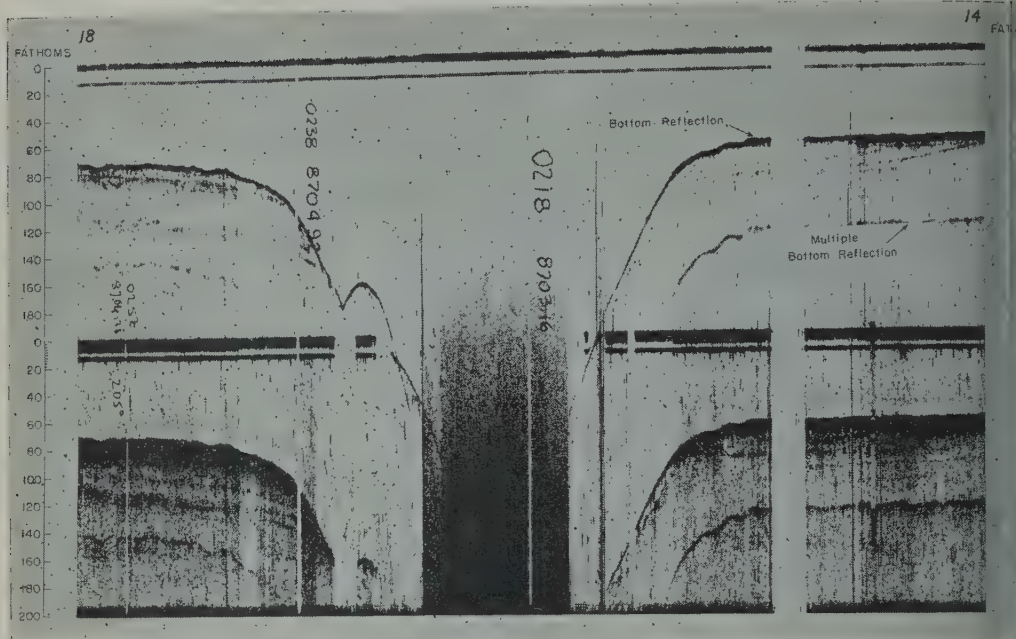


Fig. 4. SDR record, profile 18-14.

are the location of the core samples obtained during the survey.

Figures 3 to 7 show reproductions of records made on several traverses in the vicinity of the canyon. These indicate the continuity of various horizons and show more details of the structure than can be shown on the scale of the profiles in Figure 2. The lines parallel to the canyon (15-20, 17-19, 18-21, and 7-2; see Fig. 1) were particularly useful for correlating horizons between the cross-canyon profiles and for showing the attitude of the layering near the edge of the shelf. These show that, in general, deeper horizons, presumably bedding planes, dip more steeply seaward than does the sea floor. In several places there is clear evidence of an unconformity at about the 80-fm depth where the tilted beds are truncated and recovered by later deposition, e.g. profiles 15-20 and 14-23 (Fig. 5). This unconformity apparently represents the erosional surface associated with a former stand of sea level. Figures 4 and 5d show examples of benches just above the break in slope probably formed by the erosion of an overlying, presumably less competent, bed. Figure 6, a profile across the edge of the shelf, clearly shows a large pocket of sediments between depths of

about 100 fm and 140 fm lying in a step in the underlying layer. Note also the indications of reflecting horizons associated with the benches at depths of about 250 fm and 320 fm. There is little doubt that here the topography of the upper part of the continental slope is controlled by the stratification. Figure 7 shows the track from point 7 to point 2. This track is approximately perpendicular to the strike of the bedding planes, and from it the average value of dip for this area has been computed to be about  $0.5^\circ$ .

On traverses parallel to the strike of the bedding planes, the horizons are seen to outcrop on the northeast wall of the canyon with little or no change in slope. This leaves little doubt that the canyon has been cut by erosion (probably subaerial, as proposed by Veatch and Smith [1939] and others) into the older sediments of the shelf. On the southwest wall, many of the reflecting horizons, particularly the deeper ones, turn downward as they approach the wall. This downwarp is an indication that erosion has extended to a greater depth on the southwest bank of the canyon than on the northeast. This might be interpreted as confirmation of the hypothesis advanced by Veatch and Smith



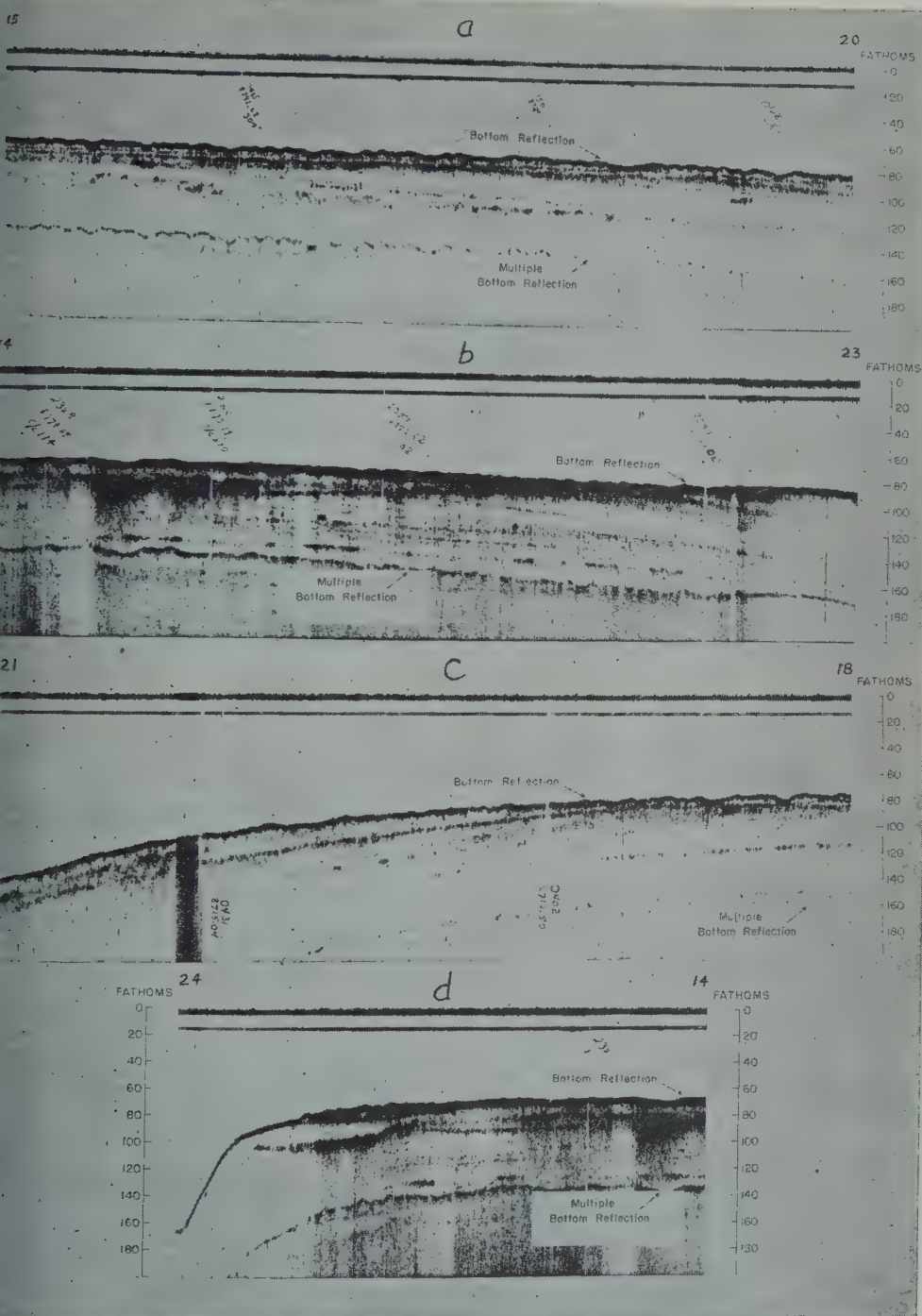


Fig. 5. SDR records, profiles 15-20, 14-23, 21-18, and 24-14.



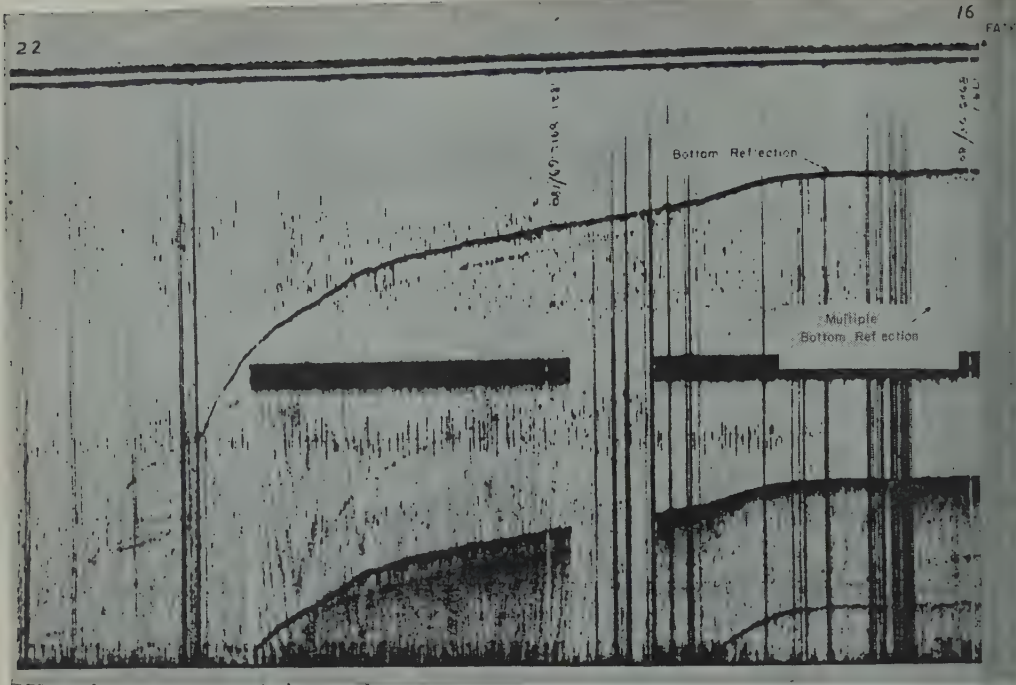


Fig. 6. SDR record, profile 22-16.

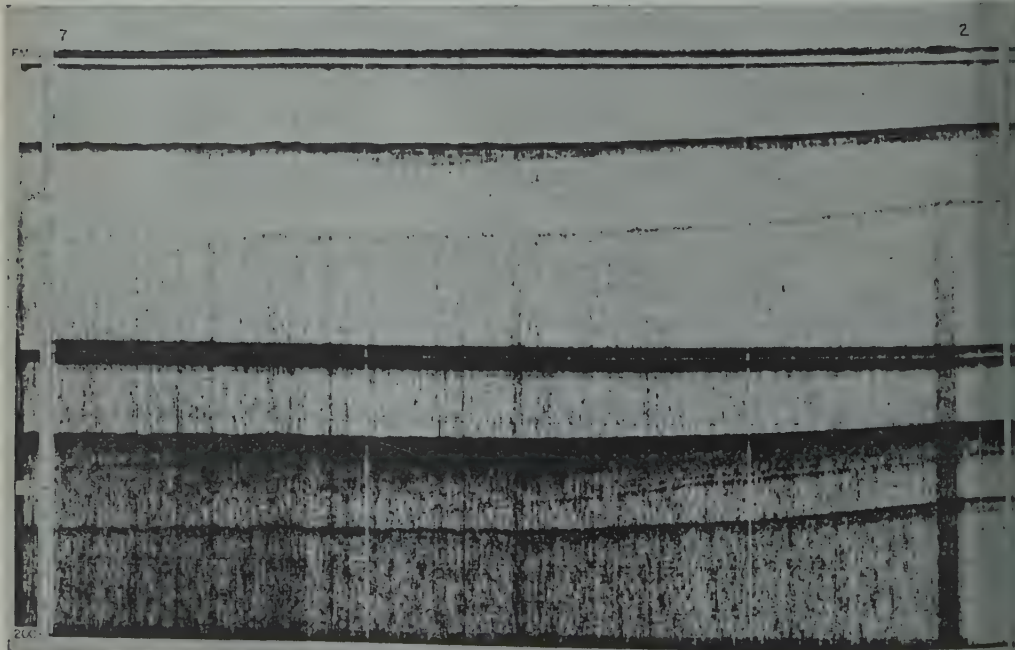


Fig. 7. SDR record, profile 7-2.

orthward migration of the Hudson Channel. It could be pointed out that, even if a sub-bottom horizon is perfectly horizontal at an angle on a slope, the echogram will show it to be somewhat conformably with the topography if there is an appreciable velocity contrast between the overlying layer and the water. The fact that many horizons on both sides of the canyon outcrop horizontally indicates that the velocity in the upper sedimentary layers is less than that in water.

The track 0-3 which crosses northwest of the head of the canyon gave inconclusive results about the possible existence of a buried channel leading to it. The track 7-2 (Fig. 7), on the other hand, definitely showed that there is a buried channel leading to the northeastern head of the canyon. The seaward-dipping beds truncated by a horizontal unconformity at a 60-fm depth, and no other sign of a channel is visible. Hence it can be concluded that, if a branch of the canyon ever had a significant channel leading to it, the bottom of the channel was no deeper than the present 60-fm depth and all the material into which it was eroded must subsequently have been removed by erosion.

Six sediment cores were taken during this survey. Their locations relative to topography are shown in Figure 2. Core 1, taken on the northern lip of the canyon near the north end, is 95 cm long and consisted mostly of fine sand and lutite. All the material in the core probably came from the upper layer. Core 2, 100 cm long, was taken on the eastern lip of the canyon near the southern limit of the area surveyed and consisted almost entirely of gray sand with the exception of a thin sand layer at the top. The material is probably entirely from the top layer. Core 3, taken in the northern portion of the survey area on the west wall of the canyon, was 585 cm long and consisted mostly of sandy lutite. It is possible that the bottom of the core reached the first sub-bottom horizon, since that part of the sample showed a sharp color contrast and an increase in amount of shell fragments. Core 4, taken in the center of the survey area on the east wall of the canyon, struck very near the point of the second sub-bottom horizon outcrops. The material in sample 4 was much coarser

than that in the other cores, containing a high percentage of sand, pebbles, and shell fragments. It is logical to suppose that this coarse layer is the reflector.

The foregoing description of the core samples is based on shipboard examination. These samples are being analyzed and will be described in detail in a supplementary report.

*Bermuda Platform, Challenger Bank, and Plantagenet Bank.* An SDR line run from St. Georges harbor, along the southern coast of Bermuda, and across Challenger and Plantagenet banks showed no indication of sub-bottom horizons. Approximately 8 feet of calcareous sand was obtained in a core from the top of Plantagenet Bank, demonstrating that some soft cover exists; but it is either too thin to be resolved by the SDR or is unstratified in the upper region.

*West Indies Arc, Virgin Islands, and Puerto Rico.* On two traverses across the island arc east of Puerto Rico, the SDR showed sub-bottom horizons in only a few spots, most of them north of the Virgin Islands. Figure 8 shows a section of SDR record from a crossing of the northern edge of the shelf northwest of Culebra. The sounding track is shown in Figure 9.

*West Atlantic Basins.* Sub-bottom reflections recorded on standard echo sounders have been observed and discussed by Hersey and Rutstein [1958], Heezen, Tharp, and Ewing [1959], [Worzel] 1959, and others. Worzel was able to obtain core samples of sufficient length to identify a prominent reflector in the southeast Pacific as a widespread layer of white ash. During the run from New York to Puerto Rico the echo sounder was operated as much of the time as possible with a short (5-msec) pulse in order

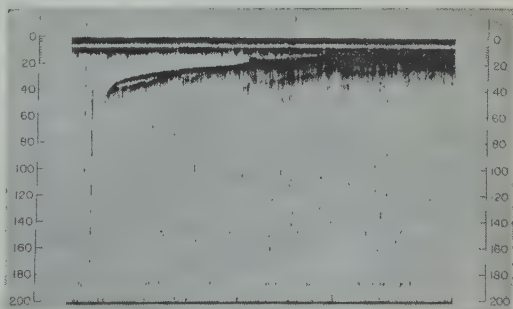


Fig. 8. Sample SDR record northwest of Culebra Island in the West Indies.

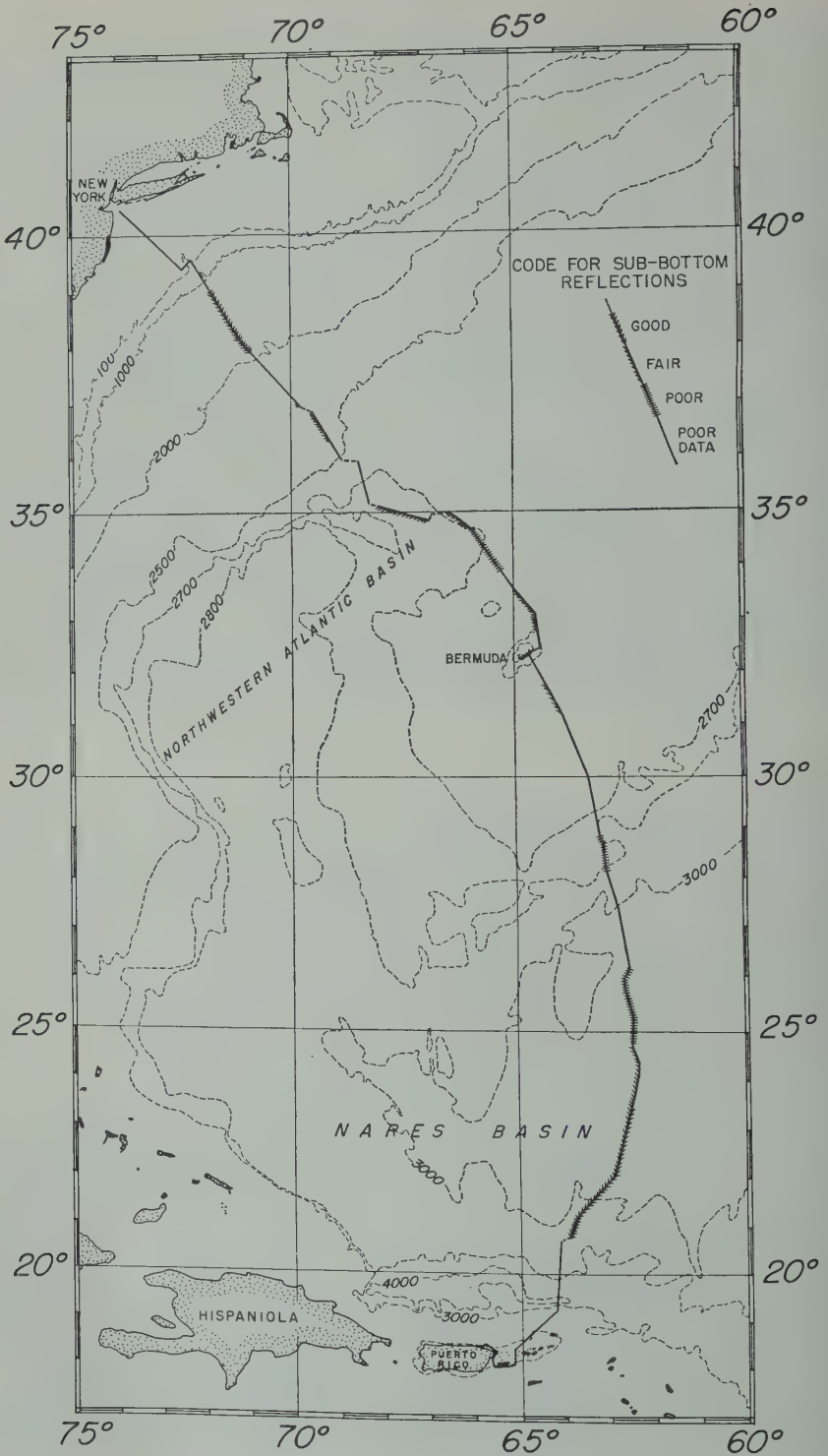
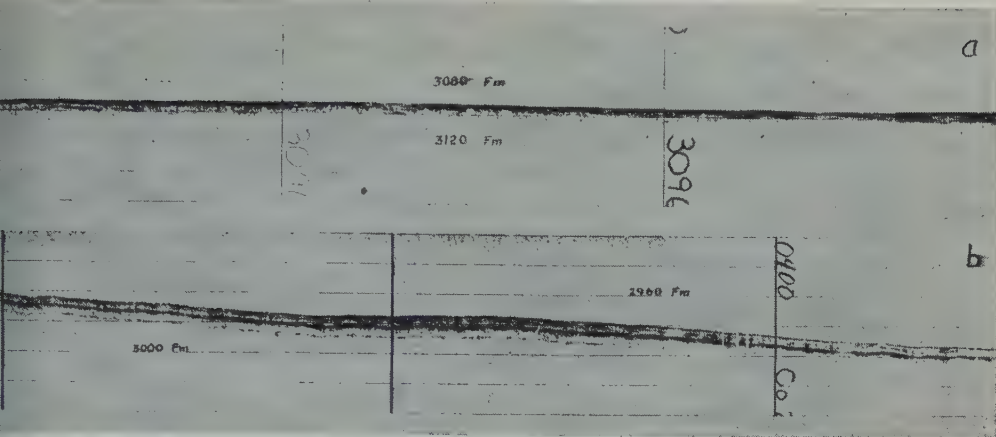


Fig. 9. Vema track from New York to San Juan showing areas of sub-bottom reflection observat





10. PDR records: *a*. Abyssal plain of Nares Basin. *b*. Outer rise of Puerto Rico Trench.

tain the maximum amount of sub-bottom reflection data. Figure 9 shows *Vema's* track areas of good, fair, and poor reflections included. Along the remainder of the track the quality of the data is such that it is not possible to decide whether sub-bottom reflections are present.

In general, sub-bottom reflections in abyssal areas are not consistent. Frequently, many reflections can be seen, but they are apt to fade or disappear in relatively short distances. Figure 10 shows a sample record from the abyssal plain in the Nares Basin. On the elevated regions, such as the Bermuda Rise and the outer rise of Puerto Rico Trench, consistent reflections are very common. The best reflections observed in the present study were on the outer rise, where one horizon, about 6 to 8 fm beneath the sea, could be traced for 30 miles. In addition, one shallower and one deeper reflector can be seen over most of the same track. Figure 10b shows 12 miles of track in this area.

A sediment core 23.6 meters in length was recovered at 2767 fm on the northern flank of the outer rise at 20°52.5'N, 64°04'W. Figure 11a shows the fathogram beginning 15 minutes before and ending 15 minutes after the core was struck. The left one of the two vertical lines indicates the time of contact. In the same figure, *b* and *c* are fathograms made while the ship was underway shortly before and shortly after stopping at the coring station. These are shown to illustrate the possible misconception that the core

was taken on an abyssal plain, since the 'drifting' record made here has somewhat the same appearance as an 'underway' record in some abyssal-plain areas.

The most persistent reflectors are at depths of about 7.5, 14, 20, and 26 meters (velocity un-

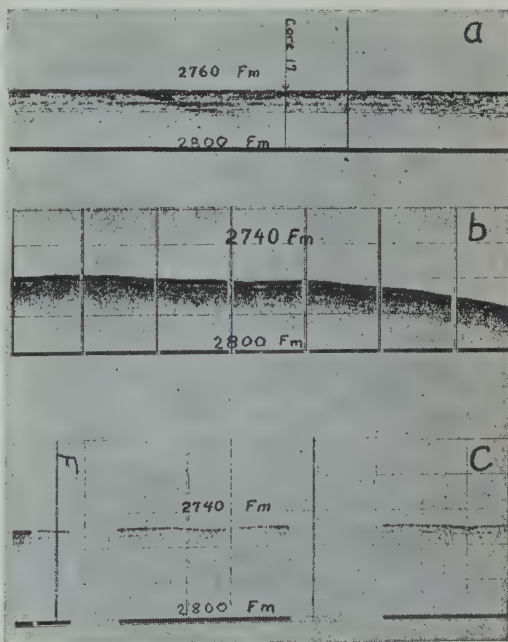


Fig. 11. PDR records at coring station on outer rise of Puerto Rico Trench: *a*. 'Drifting' record at core location. *b*. 'Underway' record before core station. *c*. 'Underway' record after core station.



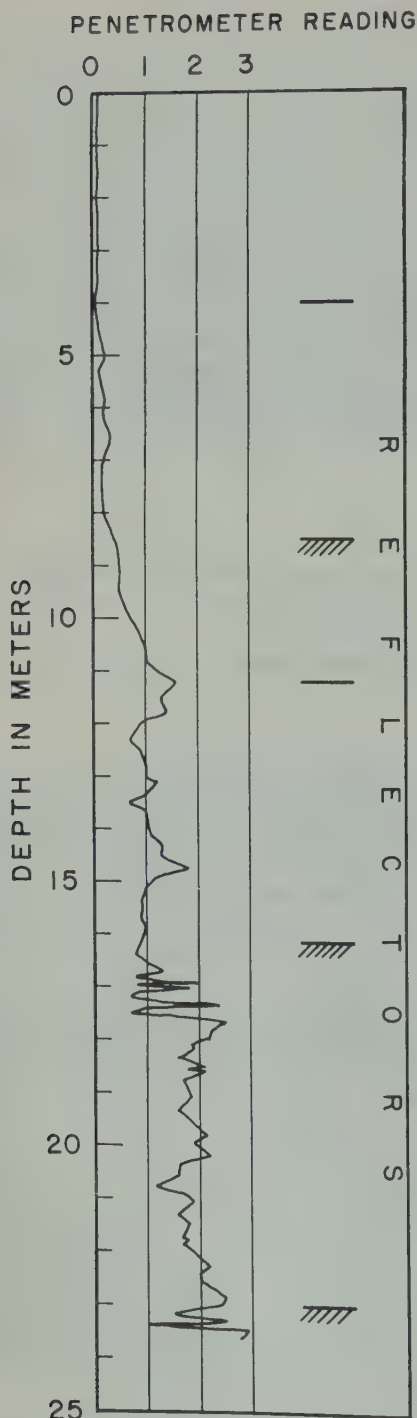


Fig. 12. Comparison of penetrometer readings and reflector depths in Core V-16-17. Strongest reflectors are underscored with hachures.

corrected). For short periods of time other reflections appear at depths of about 3.5, 10, and 40 meters. Figure 12 is a plot of penetrometer readings versus depth in the core compared with corrected reflection depths. Corrections in depth have been applied for sedimentary velocities of 1.6 km/sec at the top to 1.7 km/sec at the bottom of the core. Although velocity measurements were not made for this core, the above values are in accordance with measurements from other deep-sea cores [Stearns, Berckhemer, and Nafe, 1957]. There is a correlation between reflectors and changes in rigidity at depths of 8.5, 16, and 23 meters. The core is only slightly longer than 23 meters, so it is not known whether there is an overall increase in rigidity below this point, as at other reflector depths; but the large fluctuations in rigidity beginning here could account for the reflections. The reflector at 4 meters is a poorly resolved. If present, it probably represents the first measurable increase in rigidity which occurs between 4 and 5 meters. The reflectors at 8.5 and 11.2 meters correspond to a zone in the core where the penetrometer readings increased from about 0.2 to 1.0. In the same region, at 9.8 meters, there is a thin layer which has a higher silt content than any other part of the core and is noticeably different in texture. It is much firmer but unfortunately was not measured with the penetrometer because the core was cut very near to the layer after extrusion. In a nearby core a layer believed to be the same, gave a penetrometer reading of 5.3. Neither of the two reflections at this approximate depth of this layer coincides with it; hence it is deduced that, in spite of its distinct textural contrast with the rest of the core, it is not a good reflector. The explanation for this may lie in destructive interference, since the layer thickness ( $\sim 6-7$  cm) is approximately equal to  $\frac{1}{2}$  wavelengths of the frequency of the sounder.

In the region about 8.5 to 10 meters there is a distinct increase in the amount of foraminiferal material, and below this level the core is characterized by many alternating light-tan and reddish-brown lutite, the latter color being associated with higher carbonate content. At a depth of 16 to 17 meters, one of the strong reflections occurs, the average

trometer reading increases from about 1 to 1.9, although there is no obvious change in the core material. A careful laboratory analysis may reveal a change in porosity, density, median grain size, if not in carbonate content. These properties have all been shown to affect acoustic impedance [Sutton, Berckhemer, and Nafe, 1957].

### SUMMARY

The SDR traverses reported here constitute more than a reconnaissance of the continental shelf near the Hudson Canyon, but they have demonstrated the great possibilities of combined coring and sub-bottom reflection measurements, particularly in areas such as this where reflecting horizons can be traced to outcrops. This study has shown that the bedding planes in this area of the continental shelf dip seaward at an angle of 0.4 to 0.5°. The beds are truncated by an approximately horizontal erosional surface at a depth of 80 to 90 fm near the edge of the shelf and by a similar surface at approximately 150 fm in the area around the head of the canyon. These areas appear to correlate at least approximately with areas which would have been subjected to wave erosion during the periods when the sea stood at the levels associated with the Wisconsin and Franklin shores, respectively [Nicholls and Smith, 1939]. Both of these erosional surfaces are now completely or partially covered by more recent deposition. A good correlation of sub-bottom reflections with changes in rigidity has been found in a core sample taken from the outer ridge of the Puerto Rico Trench. At least one important change in rigidity coincides with an increase in the amount of calcium carbonate in the core. Improvements in echo-sounding equipment and techniques and in apparatus for obtaining longer core samples are leading to important advances in the study of sedimentation in the oceans. The ability to detect and trace sub-bottom horizons over fairly wide areas not only serves as a valuable guide in determining ideal positions for drilling but also eliminates uncertainties in lateral extrapolation.

**Acknowledgments.** This work was made possible through the support of the Bureau of Ships, Contract NObSR 64547, and the Office of Naval Research, Contract Nonr-266 (48), of the U. S. Navy. The assistance of the following members of the scientific crew of *Vema* is gratefully acknowledged: Stephen Chelminski, Tom DeWitt, Stefan Gereben, Don Koelsch, Xavier LePichon, Fred Rosselot, Peter Schreiber, Lee Simon, Robert Wall, and Thayer Willis.

We are indebted to Maurice Ewing for many discussions about the significance of the data reported here, particularly with regard to the identification of certain reflecting horizons as probable erosional surfaces connected with former sea levels.

### REFERENCES

- Beckmann, W. C., A. C. Roberts, and B. Luskin, Sub-bottom depth recorder, *Geophysics*, **24**, 749-760, 1959.
- Ewing, M., J. L. Worzel, N. C. Steenland, and F. Press, Geophysical investigations in the emerged and submerged Atlantic Coastal Plain, 5: Woods Hole, New York, and Cape May sections, *Bull. Geol. Soc. Am.*, **61**, 877-892, 1950.
- Hersey, J. B., and M. S. Rutstein, Reconnaissance survey of Oriente Deep (Caribbean Sea) with a precision echo sounder, *Bull. Geol. Soc. Am.*, **69**, 1297-1304, 1958.
- Heezen, B. C., M. Tharp, and M. Ewing, The floors of the ocean, 1, The North Atlantic, *Geol. Soc. Am. Spec. Paper* **65**, 1959.
- Knott, S. T., and J. B. Hersey, High resolution echo sounding techniques and their use in bathymetry: Marine Geophysics and Biology, *Deep-Sea Research*, **4**, 36-44, 1956.
- Luskin, B., B. C. Heezen, M. Ewing, and M. Landisman, Precision measurement of ocean depth, *Deep-Sea Research*, **1**, 131-140, 1954.
- Luskin, B., and A. C. Roberts, *Lamont Geol. Observatory Tech. Rept.* **6**, CU-15-55-N6, Onr 27124 Geol, 1955.
- Oliver, J. E., and C. L. Drake, Geophysical investigations in the emerged and submerged Atlantic Coastal Plain, 6: The Long Island area, *Bull. Geol. Soc. Am.*, **62**, 1287-1296, 1951.
- Sutton, G. H., H. Berckhemer, and J. E. Nafe, Physical analysis of deep sea sediments, *Geophysics*, **22**, 779-812, 1957.
- Veatch, A. C., and P. A. Smith, Atlantic submarine valleys of the United States and the Congo submarine valley, *Geol. Soc. Am. Spec. Paper* **7**, 1939.
- Worzel, J. L., Extensive deep sea sub-bottom reflections identified as white ash, *Proc. Natl. Acad. Sci. U. S.*, **45**, 349-355, 1959.

(Manuscript received May 31, 1960.)



# On the Mass Transport through the Drake Passage

FEODOR OSTAPOFF

U. S. Weather Bureau  
Washington, D. C.

**Abstract.** The mass transport through the Drake Passage has been computed geostrophically on the basis of a nonconstant reference level. The reference level has been determined on the assumption that the Antarctic Circumpolar Current behaves like an equivalent-barotropic system. For such a system a total eastward transport of  $40 \times 10^6 \text{ m}^3 \text{ sec}^{-1}$  has been calculated; a counter flow underneath amounts to  $40 \times 10^6 \text{ m}^3 \text{ sec}^{-1}$ .

The geostrophic concept has been applied to oceanic currents with various degrees of success, and the choice of a reference level for the transformation of relative dynamic topographies into absolute topographies has presented a major problem in oceanography. Dramatic results have been obtained by G. Wüst not only in the case of the Florida Current [Wüst, 1924] but also, recently, in the case of the South Atlantic sea circulation [Wüst, 1957]. Wüst's success is due to a great extent from the choice of the most probable depth where the horizontal density gradients vanish or become negligibly small. In the case of the Florida Current he used the use of current observations by J. E. Pillsbury. In the second case, he used Defant's level of no motion [Defant, 1941]. The crucial importance of the reference level when dealing with transports of properties like mass, salt, and heat has been stressed by Dietrich [1937] in connection with the dynamics of wind-driven ocean currents, by Neumann [1955].

This paper presents an alternative analysis of geostrophic observations obtained by the *Discovery* in April 1930, through the application of recent work in dynamic oceanography. For this study of mass transport through the Drake Passage portion of the Antarctic Circumpolar Current, *Discovery* stations 382 to 388 have been selected.

In order to illustrate the dynamic conditions across a section through the Drake Strait, Figure 1 shows the distribution of the density  $\sigma_t$ . The isopycnals slope strongly in the region, near the Antarctic Convergence or Polar Front commonly placed, and also near the conti-

nental shelf off Cape Horn. Except for the weaker stratification, this section reveals a density distribution similar to that found in sections of the Gulf Stream off the northeast coast of the United States [see, for example, Neu-

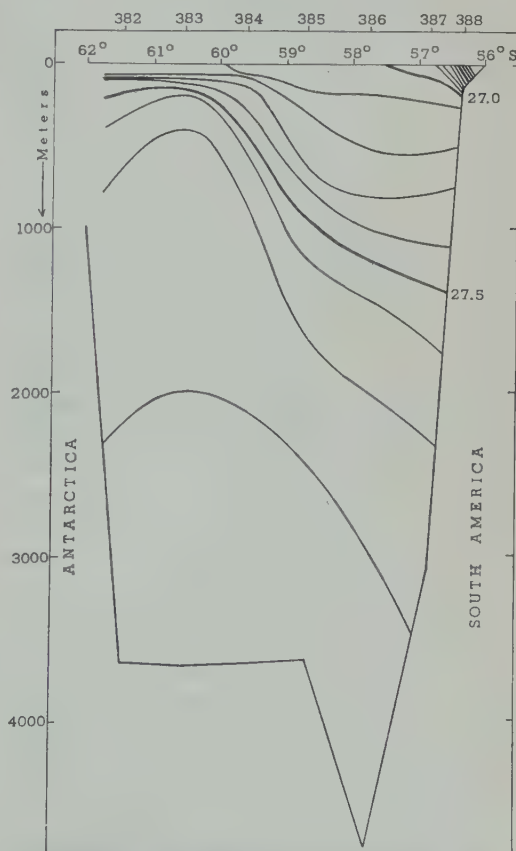


Fig. 1. The density  $\sigma_t$  distribution.



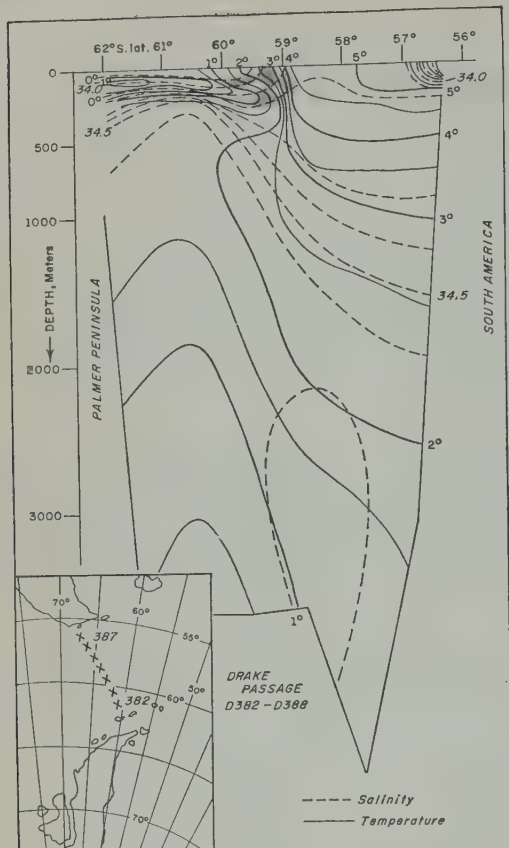


Fig. 2. The temperature-salinity distribution. Insert map shows the geographical location of the cross section.

mann, 1956]. In Figure 2, where the temperature distribution is superimposed on the salinity distribution, the difference in the thermohaline structure can be seen between the isopycnal slopes between 60°S and 59°S and those near the South American shelf. In the first case the isohalines cross the isotherms with a large angle forming a solenoidal field which is indicated in Figure 2 by hatching. A similar thermohaline distribution also is found at the left-hand side of the Gulf Stream (for example, *Atlantis* stations 5295 to 5301). Near Cape Horn, however, the isotherms parallel, more or less, the isohalines, suggesting that the density distribution is determined mainly by continental influences [Krauss, 1958].

In order to compute the geostrophic mass transport a reference level must be assumed or

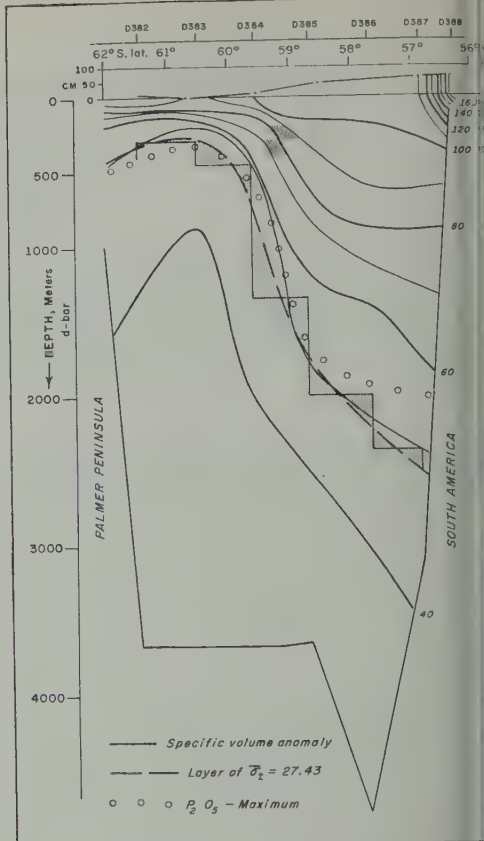
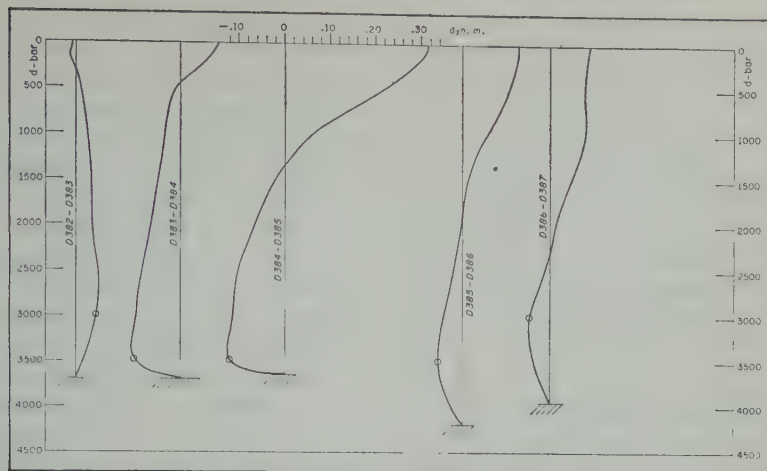


Fig. 3. The specific volume anomaly distribution. Isobaric-isosteric solenoids are indicated by hatched areas. The sea surface slopes are indicated on an exaggerated scale.

determined. Previous investigators assumed 3000 decibar or 3500 decibar<sup>1</sup> surface as reference level [Clowes, 1933; Kort, 1959]. Defant [1941] constructed a topography of reference surface for the Atlantic Ocean [Neumann, 1955]; however, his study is confined to the area between 50°N and 10°S, leaving us without a clue as to how this topography should continue through the Antarctic Circumpolar Current and approaching the arctic Continental Shelf. Furthermore, Defant's method, when applied to *Discovery* stations 382 to 387, leads to ambiguous results, especially when no other support is available. Comparisons of other oceanographic sections around Antarctica exhibit in each case a

<sup>1</sup> 1 decibar numerically approximates 1 meter.



Curves of dynamic depth differences between two stations as function of depth. All curves are drawn on the same scale. Circles indicate the depth of the observations.

similar to the density distribution shown in Figure 1. An attempt was made to apply to the Arctic Circumpolar Current the findings by *ann* [1956] and *Martineau* [1957], which were obtained in the Gulf Stream, and to use them as a guide for determining the probable shape of the reference surface. Both authors stated that, if Defant's reference level is chosen, the mean density from the sea surface to the level of no motion varies only slightly in the cross-stream direction. For a constant mean density the system is equivalent to a two-layer model with an inclined interface.

At every station 387 has been used for the determination of the mean density value, which can be used for the equivalent-barotropic model system. The mean density

$$\bar{\sigma}_t = \frac{1}{D} \int_0^D \sigma_t dz$$

has been obtained by means of planimetry in such a way that the depth  $D$  of the reference level smoothly joins Defant's topography for the Atlantic Ocean, which for this purpose has been slightly extrapolated into the Drake Passage. The value  $\bar{\sigma}_t = 27.43$  has been found for the layer between the sea surface and 2500 meters. The graphical integration was performed at the other stations of the section until the same mean density value was obtained. Assuming a value of 27.43 for the mean density of the model system, an inclined interface is found,

which is identified with the reference level. Thus, the reference level slopes downward to the north from approximately 350 meters at 61°S to 2500 meters at 57°S, as can be seen in Figure 3, where the distribution of the specific volume anomaly is also plotted; the same general shape is followed by the oxygen minimum and the phosphate maximum, the latter being indicated in Figure 3 by circles.

The orientation of the slope of the reference level is in agreement with the results which follow from a two-layer model. The interface between the two water bodies with different mean densities must be inclined as is shown in Figure 3, which follows from Margules' equation. Some further support is given to the general shape of the reference level in this region by the investigations of *Krauss* [1958, pp. 346-352]. He computed the zonal components of the geostrophic current on the basis of a density distribution which results from the temperature distribution due to absorption of solar energy, and the salinity distribution due to the influence of evaporation and precipitation. In *Krauss's* paper the 0.5 cm sec<sup>-1</sup> isotach slopes upward with increasing latitude.

However, our results are in disagreement with the results of *Mamaev* [1957]. Generally, the depth of *Mamaev's* reference level increases continuously toward Antarctica, reaching approximately 3000 meters off the Palmer Peninsula; according to our analysis, however, the depth

should decrease through the region of the Antarctic Circumpolar Current and then remain nearly constant or possibly increase slightly toward the Antarctic Continental Shelf. Mammaev's method of computing the depth of the reference level is based on very restrictive assumptions, such as a linear vertical density and velocity profile everywhere, which cannot be supported by observations. On the contrary, the observed density profiles south of the Circumpolar Current are strongly nonlinear, with most of the changes occurring in the upper 200 meters.

On the basis of this reference level the total mass transport has been computed. In Figure 4 the dynamic depth differences as function of depth are drawn for all station pairs outside the continental shelf. These curves also represent the geostrophic velocity profile normal to the cross section down to the depth indicated by a circle in the figure; from this point the curve has been continued to the bottom so that there the velocity vanishes. These curves have been used for the mass transport calculations for the total layers. From Figure 4 it can be seen that for the two curves at the right-hand side of the diagram the choice of the level of zero-geostrophic velocity is quite consistent with the level which one would obtain by applying Defant's method. If for the two curves at the left-hand side of Figure 4 (382-383 and 383-384) the zero level had been chosen approximately 200 meters deeper than is shown in the figure, then that level also could be interpreted according to Defant's concept as a possible reference level. Assuming that the reference level is to be located slightly deeper than would follow from the equivalent-barotropic analysis, the mean density  $\bar{\sigma}_t$  would increase south of the Circumpolar Current. This small increase of density in cross-stream direction also has been found persistently in the analyzed Gulf Stream sections by Neumann [1956] and Martineau [1957], thus indicating a possible departure from the equivalent-barotropic state in the region of anticyclonic shear.

Table 1 comprises the results of the computations and presents the velocity components perpendicular to the cross section. Negative values indicate components directed toward the west.

The largest velocities are found between sta-

tions 384 and 385 (see Fig. 2) which are in the region of steepest slope of the isopycnals. If the surface layer moves with a velocity of more than 20 cm sec<sup>-1</sup> toward the east whereas the westward motion with maximum velocities more than 8 cm sec<sup>-1</sup> is found at depths between 3000 to 3500 meters, some few hundred meters above the ocean floor. These values represent averages over a distance of 118 km. In general the velocities at the surface decrease from the maximum toward both continents, except on the shelf near the South American coast. The computations terminate at station 387, leaving out the strong but very shallow coastal current between station 387 and the shelf station 388. On the Antarctic side of the section between stations 382 and 383 a weak current moving toward the west develops in the surface layer.

TABLE 1. The Geostrophic Velocity Components Normal to the Cross Section

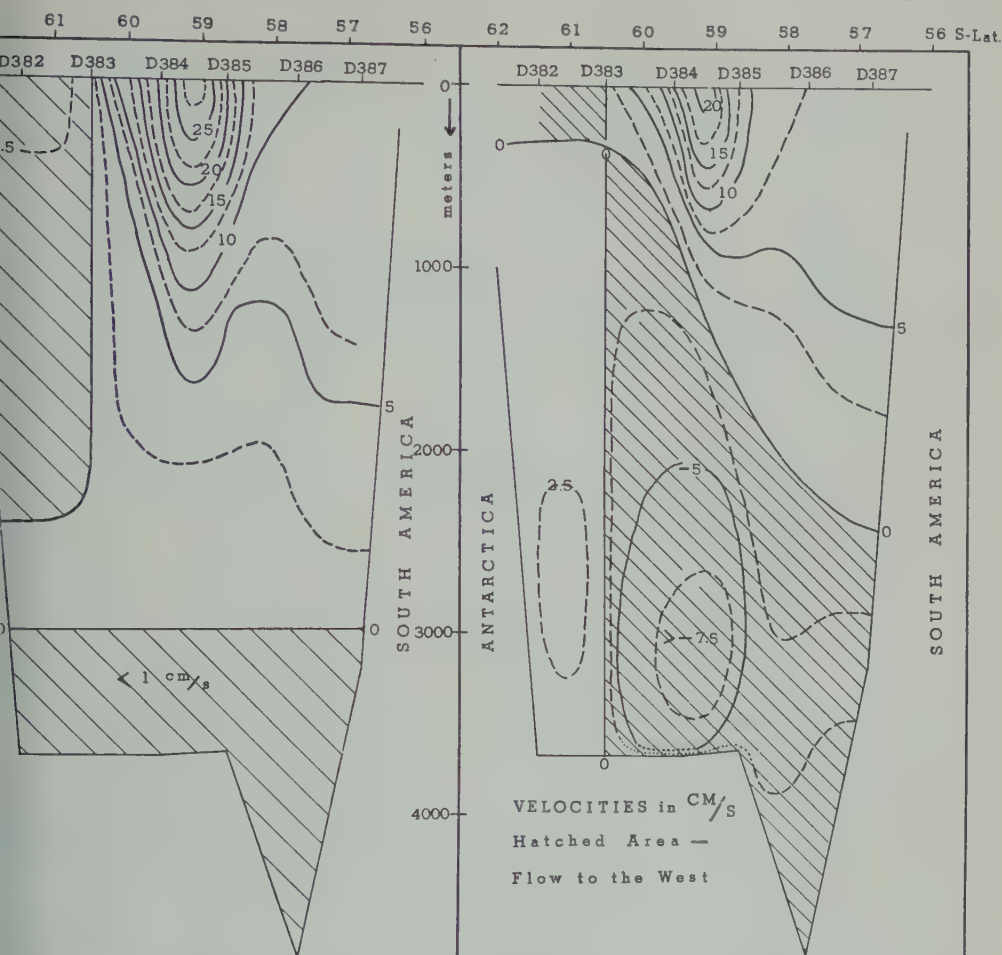
| Stations                    | 382-383          | 383-384 | 384-385 | 385-386 | 386-387 |
|-----------------------------|------------------|---------|---------|---------|---------|
| Distance, km                | 118.0            | 117.8   | 118.0   | 118.0   | 118.0   |
| Mean Latitude               | 61°00'           | 60°02'  | 59°09'  | 58°13'  | 57°18'  |
| Reference Level, m or d-bar | 300              | 450     | 1350    | 2000    | 2500    |
| Depth, m                    | Velocity, cm/sec |         |         |         |         |
| 0                           | -0.5             | 5.7     | 21.2    | 8.7     |         |
| 200                         | -0.6             | 4.0     | 19.4    | 8.3     |         |
| 400                         | 0.5              | 0.6     | 16.4    | 7.5     |         |
| 600                         | 0.9              | -1.3    | 12.4    | 6.7     |         |
| 800                         | 1.1              | -1.8    | 7.8     | 5.5     |         |
| 1000                        | 1.4              | -2.0    | 4.2     | 4.0     |         |
| 1200                        | 1.9              | -2.5    | 1.7     | 2.6     |         |
| 1400                        | 2.2              | -3.0    | -0.5    | 1.6     |         |
| 1600                        | 2.4              | -3.4    | -2.5    | 0.9     |         |
| 1800                        | 2.4              | -3.7    | -3.7    | 0.5     |         |
| 2000                        | 2.3              | -4.1    | -4.8    | 0.0     |         |
| 2200                        | 2.5              | -4.7    | -5.9    | -0.4    |         |
| 2400                        | 3.0              | -5.4    | -6.8    | -0.8    |         |
| 2600                        | 3.2              | -5.9    | -7.4    | -1.2    |         |
| 2800                        | 3.1              | -6.2    | -7.7    | -1.8    |         |
| 3000                        | 3.0              | -6.4    | -7.9    | -2.4    |         |
| 3200                        | (2.6)*           | -6.6    | -8.2    | -3.0    |         |
| 3400                        | (1.9)            | -6.8    | -8.5    | -3.4    |         |
| 3600                        | (0.8)            | (-5.8)  | (-5.2)  | (-3.5)  |         |
| 3800                        | (0.0)            | ( 0.0)  | ( 0.0)  | (-3.2)  |         |
| 4000                        |                  |         |         | (-2.3)  |         |
| 4200                        |                  |         |         | ( 0.0)  |         |

\* Extrapolated values in parentheses.



## DRAKE PASSAGE

## DISCOVERY 382-387



5. The mean velocity field through the Drake Passage. (a) The velocity distribution relative to the 27.43-decibar level. (b) The velocity field in an equivalent-barotropic flow of a mean density  $\bar{\sigma}_t = 27.43$ .

Figure 3 the sea surface slope has been indicated by plotting it on exaggerated scale. Between stations 383 and 387 a sea-level difference of approximately 60 cm, sloping downward toward the south, results from our analysis, as compared with 90 cm when the 3000-decibar level is the reference level. If the latter reference level is used the surface velocity increases by 21 cm sec<sup>-1</sup> between stations 384 and 385, compared with 21 cm sec<sup>-1</sup> in our case. This is a noticeable difference in terms of velocity, but it is of far-reaching consequences in terms of mass transport.

Figure 5 represents an attempt to illustrate the mean velocity field through the Drake Passage.

This representation is merely schematic because of the uncertainty in placing isotachs based on space mean values with artificial discontinuities at the station verticals resulting from the replacement of derivatives by finite differences. Comparison of the velocity field relative to a 3000-decibar level (Fig. 5a), which has been assumed generally, with that for our two-layer model (Fig. 5b) reveals certain distinctive differences:

(a) The surface westward current near the Antarctic continental shelf is, in our analysis (Fig. 5b), of only relatively shallow nature, whereas the choice of a 3000-decibar reference



TABLE 2. The Mass Transport through the Drake Passage under the Assumption of a Two-Layer System (in million tons sec<sup>-1</sup>)

| Stations | Reference Level, decibar | Above the Reference Level | Below the Reference Level |
|----------|--------------------------|---------------------------|---------------------------|
| 382-383  | 300                      | -0.2*                     | +8.2*                     |
| 383-384  | 450                      | +1.3                      | -16.3                     |
| 384-385  | 1350                     | +16.6                     | -16.2                     |
| 385-386  | 2000                     | +10.6                     | -5.2                      |
| 386-387  | 2350                     | +11.2                     | -3.6                      |

\* Positive values mean eastward transport, negative values westward transport.

level (Fig. 5a) imposes a triple regime on the whole water column, extending the upper westward-moving layer to a depth of 2400 meters, below which a layer 600 meters thick moves eastward and at depths below 3000 meter the water turns westward again.

(b) In the portion of the cross section, which at the surface is distinguished by a current maximum (between stations 384 and 385), a counter current below 3000 meters is found in both cases; however, in our representation (Fig. 5b) a well-pronounced current maximum is found in the 3200- to 3400-meter layer, which maximum is almost nonexistent in Figure 5a.

The mass transport amounts are summarized in Table 2 and in Figure 6, which shows schematically the cross section through the Drake Passage. The reference level has been indicated as a continuous band separating the two current systems. The Antarctic Circumpolar Current flows eastward and carries approximately  $40 \times 10^6 \text{ m}^3 \text{ sec}^{-1}$ . As mentioned above, this figure does not include water transported by the shallow current near or on the South American Shelf. Under the Circumpolar Current some  $40 \times 10^6 \text{ m}^3 \text{ sec}^{-1}$  return to the Pacific Ocean, resulting in a strong counter current. Near the Palmer Peninsula and below the very weak ( $2 \times 10^6 \text{ m}^3 \text{ sec}^{-1}$ ) westward current surface a relatively strong ( $8 \times 10^6 \text{ m}^3 \text{ sec}^{-1}$ ) eastward movement of water mass<sup>2</sup> results.

It should be stressed that there is a priori no

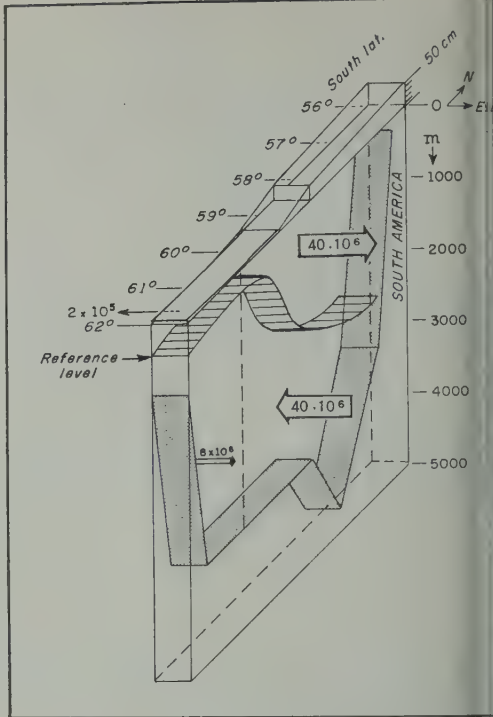


Fig. 6. Schematic representation of the mass transport in tons sec<sup>-1</sup>.

reason to assume that in nature a current system should fulfill exactly the equivalent-barotropic conditions. However, it is quite conceivable that current systems similar to the one under discussion approach this state within the limits set by other forces not considered in the analysis, e.g. the action of the wind stresses on the sea surface. In the light of these remarks it is possible that the transport figures presented in this study represent only minimum values for the Antarctic Circumpolar Current.

As with all geostrophic approximations in oceanography, these transport figures depend entirely on the choice of the reference level. Because *Sverdrup, Johnson, and Fleming* [1942] estimated the amount of mass transport in the Antarctic Circumpolar Current relative to the 3000-decibar level, it is not surprising that they found an eastward displacement of  $90 \times 10^6 \text{ m}^3 \text{ sec}^{-1}$ . *Clowes* [1933] computed a total transport through the Drake Passage of  $110 \times 10^6 \text{ m}^3 \text{ sec}^{-1}$  above the 3500-decibar surface, whereas *Koeller* [1959] obtained  $134 \times 10^6 \text{ m}^3 \text{ sec}^{-1}$  with respect

<sup>2</sup> Numerically the units of volume transport ( $\text{m}^3 \text{ sec}^{-1}$ ) closely approximate the units of mass transport (tons  $\text{sec}^{-1}$ ) and are interchangeable.

TABLE 3. The Mass Transports, Assuming for the Upper Layer that  $\bar{\sigma}_t = 27.38$  and 27.48 (in million tons sec<sup>-1</sup>)

| Stations                 | $\bar{\sigma}_t = 27.38$ |                        |             |                        | $\bar{\sigma}_t = 27.48$ |                        |             |                        |
|--------------------------|--------------------------|------------------------|-------------|------------------------|--------------------------|------------------------|-------------|------------------------|
|                          | Upper Layer              |                        | Lower Layer |                        | Upper Layer              |                        | Lower Layer |                        |
|                          | Thickness                | Transport              | Thickness   | Transport              | Thickness                | Transport              | Thickness   | Transport              |
|                          | meters                   | 10 <sup>6</sup> tons/s | meters      | 10 <sup>6</sup> tons/s | meters                   | 10 <sup>6</sup> tons/s | meters      | 10 <sup>6</sup> tons/s |
| 383                      | 0- 240                   | -0.07*                 | 240-3700    | +9.5*                  | 0- 380                   | -0.3*                  | 380-3700    | +6.7*                  |
| 384                      | 0- 380                   | +1.2                   | 380-3700    | -19.9                  | 0- 580                   | +2.4                   | 580-3700    | -11.6                  |
| 385                      | 0-1050                   | +12.4                  | 1050-3650   | -25.8                  | 0-1650                   | +21.3                  | 1650-3650   | -9.7                   |
| 386                      | 0-1750                   | +8.6                   | 1750-4200   | -6.8                   | 0-2460                   | +12.3                  | 2460-4200   | -3.2                   |
| 387                      | 0-2000                   | +8.2                   | 2000-3950   | -5.7                   | 0-2760                   | +17.3                  | 2760-3950   | -1.1                   |
| Between stations 383-387 |                          | +30.4                  |             | -58.2                  |                          | +53.3                  |             | -25.6                  |

Positive values mean eastward transport, negative values westward transport.

the 3000-decibar level and  $165 \times 10^6 \text{ m}^3 \text{ sec}^{-1}$  with respect to the bottom as reference level. However, Clowes concluded on the basis of his particular choice of reference level that the Antarctic Water in the Atlantic sector of the Antarctic Ocean must be strongly influenced by mixing of Pacific Deep Water and that 'no water of Pacific origin passes into the Pacific.' A transport from the Pacific Ocean into the Atlantic is less than that calculated by Clowes, or a current reversal in the deep layers, would support Wüst's [1928] deductions about the isolation of the Deep Water from the thermohaline circulation scheme, which indicates a North Atlantic influence.

A complete justification of our assumption of a two-layer model cannot be given yet, although the density cross section lends strong support. However, other characteristics of the potential barotropic flow [Neumann, 1956] are not indicated in the observations so far available from the Antarctic Circumpolar Current. One important inference from the theory is that the surface velocity of the system equals the mean velocity of the whole layer. As a consequence of this relationship the vertical profile of the horizontal velocity must show an intermediate maximum. There may be two principal reasons why this has not been observed in the Antarctic Circumpolar Current, the first being the wide distribution of stations, which smooths out the detailed structure of the velocity field;

the second reason may lie in the assumption that the velocity vector does not veer with depth. Thus, any change of current direction with depth may mask a possibly existent intermediate velocity maximum.

A few words may be added with respect to the sensitivity of the choice of  $\bar{\sigma}_t$  on the transport amounts. If the limits of  $\pm 0.05$  are placed on  $\bar{\sigma}_t$ , i.e.  $\bar{\sigma}_t = 27.43 \pm 0.05$ , then the computations for the lower and upper limit lead to the following figures (Table 3). These calculations show that the system is quite responsive to such changes in  $\bar{\sigma}_t$ ; for example, a decrease in  $\bar{\sigma}_t$  by .05 units reduces the transport in the upper layer by 25 per cent and correspondingly increases the transport in the lower layer by approximately 45 per cent. Corresponding changes in the opposite sense are obtained by using the upper limit of  $\bar{\sigma}_t$ . But, in any case, a counter current is found under the Antarctic Circumpolar Current unless the reference level is assumed to be the bottom.

**Conclusion.** On the assumption that the Antarctic Circumpolar Current can be treated as a two-layer model, a reference level has been found which would result if the Circumpolar Current assumes dynamically a near equivalent-barotropic state. Density sections through the current seem to lend support to this approach, as indicated by Neumann [1952], who wrote, 'There is some evidence that the dynamics of the Gulf Stream and of the Antarctic Circum-

polar are not essentially different, when considered as three dimensional motions. . . .<sup>3</sup> Furthermore, Wüst [1935] wrote, 'It is probable that in the Drake Passage the thermo-haline structure is related to the current field—similar as in the Straits of Florida, i.e., that the isohalines and isotherms are parallel to the isotachs and that the reference level coincides with a boundary of a water mass.'<sup>3</sup> In accepting this point of view the analysis of *Discovery* section 382 to 387 shows that the depth of the reference level decreases southward, contrary to the results obtained by O. I. Mamaev. On the basis of this reference level the transport in the Antarctic Circumpolar Current through the Drake Passage is about  $40 \times 10^6 \text{ m}^3 \text{ sec}^{-1}$ , and the counter flow under the Circumpolar Current also is about  $40 \times 10^6 \text{ m}^3 \text{ sec}^{-1}$ .

Modern instrumentation provides us with the means to observe currents directly at great depth. Since, as has been pointed out, the successful application of the geostrophic concept to oceanic circulation problems depends on the assumed reference level, it seems mandatory, in addition to making standard hydrographic observations, to use current indicators like the Swallow buoy at great depth.

*Acknowledgment.* I wish to thank Mr. Morton J. Rubin for reading the manuscript and revising the English.

#### REFERENCES

- Clowes, A. J., Influence of the Pacific on the circulation in the southwest Atlantic Ocean, *Nature*, 131, 189–191, 1933.  
 Defant, A., Die absolute Topographie des physikalischen Meeresniveaus und der Druckflächen sowie die Wasserbewegungen im Atlantischen Ozean, *Wiss. Erg. Deut. Atlant. Exped. "Meteor"*, 1925–27, 6(2), 191–260, 1941.  
 Dietrich, G., Die "dynamische Bezugsfläche", ein

Gegenwartsproblem der dynamischen Ozeanographie, *Ann. Hydrog. marit. Meteorol.*, 65, 519, 1937.

- Kort, V. G., New results on the transport of arctic waters, *Inform. Bull. Sov. Antark. Exp.*, 9, 31–34, 1959 (Russian).  
 Krauss, W., Untersuchungen ueber die mittl. hydrographischen Verhaeltnisse an der Meeroberflaeche des noerdlichen Nordatlantischen Ozeans, *Wiss. Erg. Deut. Atlant. Exped. "Meteor"*, 1925–27, 5(3), 251–410, 1958.  
 Mamaev, O. I., On the question of zero dynamic height and its topography in the Southern Ocean, *Dokl. Akad. Nauk SSSR*, 117(5), 808, 1957. Also in: Vertical turbulence in the sea, the surface of no motion, *Intern. Oceanogr. Congr., Preprints*, Am. Assoc. Advance Sci., 413, 1959.  
 Martineau, D. P., The Gulf Stream as an equivalent-barotropic system, Ph.D. Thesis, New York University, 84 pp., 1957.  
 Neumann, G., Some problems concerning the dynamics of the Gulf Stream, *Trans. N. Y. Acad. Sci.*, 14, 283–291, 1952.  
 Neumann, G., On the dynamics of wind-driven ocean currents, *Meteorol. Papers*, 2(4), York Univ. Press, 33 pp., 1955.  
 Neumann, G., Zum Problem der "Dynamischen Bezugsfläche", insbesondere im Golfstromgebiet, *Deut. Hydrog. Z.*, 9(2), 66–78, 1956.  
 Sverdrup, H. U., M. W. Johnson, and R. H. Fleming, *The Oceans*, Prentice-Hall, Inc., New York, 1087 pp., 1942.  
 Wüst, G., Florida- und Antillenstrom, *Veröff. Inst. Meereskunde Berlin, Neue Folge Geogr.-naturwiss. Reihe*, Heft 12, 48 pp., 1928.  
 Wüst, G., Der Ursprung der Atlantischen Tiefwasser, *Gesellsch. f. Erdk., Zeitsch., Jubilaumsband*, 506–534, 1928.  
 Wüst, G., Die Stratosphäre, *Wiss. Erg. Deut. Atlant. Exped. "Meteor"*, 1925–27, 6(1), 2, 131, 1935.  
 Wüst, G., Stromgeschwindigkeiten und Stromeinheiten in den Tiefen des Atlantischen Ozeans unter besonderer Berücksichtigung des Tiefenwasser und Bodenwassers, *Wiss. Erg. Deut. Atlant. Exped. "Meteor"*, 1925–27, 6(2), 6. Lief., 420, 1957.

(Manuscript received March 21, 1960; revised June 9, 1960.)

<sup>3</sup> Translation by the author.



## Distribution Pattern of Rainfall in the Leeward Koolau Mountains, Oahu, Hawaii

JOHN F. MINK

*U. S. Geological Survey  
Honolulu, Hawaii*

**Abstract.** An analysis of measurements of rainfall made during 1957 in a small drainage basin on the leeward slope of the Koolau Range of Oahu shows that the rainfall at each gage relates according to a geometric progression with the distance of the gage from the summit. The rainfall increases sharply in a leeward direction from the summit to a point about 2500 feet away; thereafter it decreases at a rate more gradual than that of the initial increase. Variation in orographic rainfall with distance from the crest of the range is responsible for the pattern of distribution of rainfall on the leeward slope. Rainfall from cyclonic storms has little effect on the pattern because of its generally uniform distribution. Gages in the bottom of the drainage basin consistently catch more rainfall than gages on the adjacent ridge.

**Introduction.** One of the principal aspects of investigation of the water resources of eastern and central Oahu now being made by the U. S. Geological Survey concerns the problem of the rain falling on the leeward slope of the Koolau Range that ultimately reaches the ground-water bodies. This region, where the annual rainfall may reach 300 inches, is the chief recharge area for the great ground-water bodies that underlie the southern central parts of Oahu. Several hundred rain gages are distributed over Oahu, which has an area of 604 square miles, but very few gages have been or are maintained in the rugged mountain areas. Because of the importance of the water-budget method in the determination of the dependable yield of the ground-water basins, knowledge of the rainfall pattern in a characteristic mountain drainage area was warranted.

**Description of area and location of gages.** The drainage basin of Kipapa Stream in the central part of the leeward Koolau Range was selected for study. This range consists of the eroded remnants of a large shield volcano composed of basaltic lava. The backbone of the range extends about 25 miles north-northwest along the long axis of the island. On the leeward, or southwest, side of the range, are a number of relatively small drainage basins which typically comprise deep gulches between sharp parallel ridges. At the headwaters of the gulches the

sharp ridges merge into the backbone of the Koolau Range. The gulch of Kipapa Stream is representative of these small drainage basins. The part of the Kipapa basin covered in this study drains an area of 4.3 square miles and has a length of about 24,500 feet and a width ranging from 2000 to 7000 feet (Fig. 1).

Seventeen rain gages were installed in the drainage basin at points shown in Figure 1. Of these, fourteen (gages 1 to 14) are distributed along the crest of the ridge (Kipapa Ridge) forming the southern boundary of the basin. Extremely rugged topography and dense vegetation restricted the number of gages in the bottom of the gulch, where three (gages 30, 33, and 34) are installed in the lower third of the area. Table 1 gives the distance of each gage from gage 1, measured along a line bearing N. 74.5° E. drawn between gage 1 and gage 14. On the ridge, the gages are spaced 1850 to 3650 feet apart, except near the summit where the interval is much smaller (Figs. 1 and 2). The gages in the gulch were placed so that they formed pairs with gages on the ridge at similar distances from gage 1.

The profile in Figure 2 shows the topographic nature of the ridge on which gages 1 to 14 are located. The lowest gage on the ridge is at 1150 feet above sea level and the highest sits on a small peak at 2785 feet. From gage 1 to gage 10, which lies 2500 feet leeward of the summit, the gradient of the ridge is about 5 per cent. Near



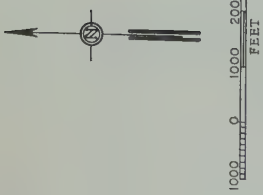


TABLE 1. Location of Rain Gages

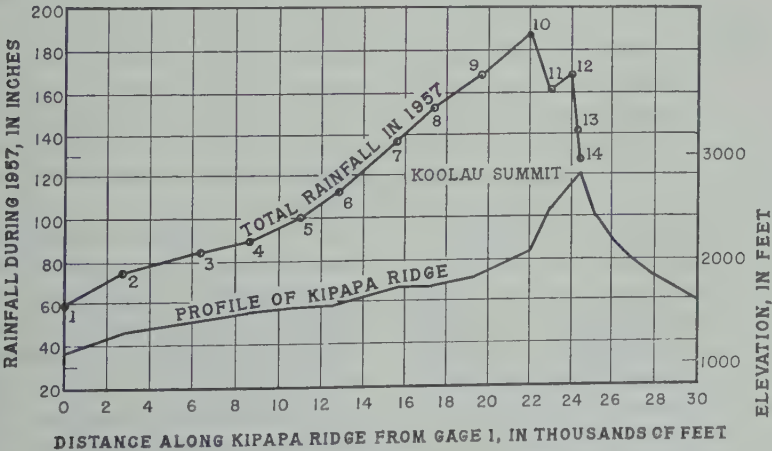
|    | Elevation above<br>Sea Level, ft | Distance from<br>Gage 1*, ft |
|----|----------------------------------|------------------------------|
| 1  | 1,150                            | 0                            |
| 2  | 1,295                            | 2,750                        |
| 3  | 1,420                            | 6,400                        |
| 4  | 1,520                            | 8,650                        |
| 5  | 1,560                            | 11,050                       |
| 6  | 1,575                            | 12,900                       |
| 7  | 1,720                            | 15,700                       |
| 8  | 1,730                            | 17,500                       |
| 9  | 1,850                            | 19,725                       |
| 10 | 2,050                            | 22,125                       |
| 11 | 2,400                            | 23,000                       |
| 12 | 2,700                            | 24,100                       |
| 13 | 2,725                            | 24,400                       |
| 14 | 2,785                            | 24,500                       |
| 30 | 720                              | 400                          |
| 33 | 870                              | 6,500                        |
| 34 | 960                              | 8,800                        |

Distances taken at intersections of normals from each gage to bearing N. 74.5° E. which connects gage 1 and ridge gage 14. Maximum distance from this bearing is 1900 feet.

On the gradient abruptly increases to about 3 per cent over the remaining distance. The steep slope of the Koolau Range is even steeper, the crest of the range is therefore sharp and narrow. The change in slope on the leeward side and the precipitous nature of the windward side of the summit ridge have marked effects on the rainfall pattern.

In the bottom of the gulch the gradient is gentle, approximately 3 per cent, up to the break in slope near gage 10. Above the break, the gulch rises in a series of steep cascades to the crest of the range.

*Description of rain gages.* The rain gages used in the study are aluminum cylinders, 3 inches in diameter and 60 inches long, made from irrigation pipe welded tight at one end. The tops of the cylinders are open and horizontally oriented, and the catch is assumed to be a direct measure of the amount of rainfall. Measurements used here were made at least once a month, on the same day of each month or on the day immediately after. Evaporation from the gages was not considered a problem because of the depth of the cylinders and the high rainfall common to the area; however, a few drops of light oil were put in the gages each month to retard any possible evaporation. A comparison of the catch of an aluminum cylinder with that of an 8-inch U. S. Weather Bureau nonrecording gage was obtained at gage 30, where a Weather Bureau gage was placed within a few feet of the aluminum cylinder. Figure 3 is a plot of the monthly rainfall measured in the aluminum cylinder against the catch of the Weather Bureau gage. The catch in the aluminum cylinder was, on the average, 0.5 inch less than that in the Weather Bureau gage at both high and low amounts of rainfall. This would be a significant difference for very low rainfalls, but it is negligible.



Profiles of the crest of Kipapa Ridge and the rainfall on the ridge during 1957. Numbers refer to ridge gages.

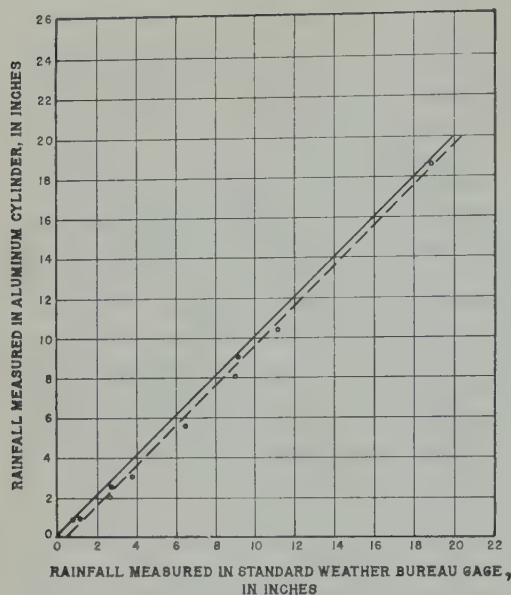


Fig. 3. Comparison of rainfall caught at monthly intervals in the 3-inch aluminum cylinder at gage 30 with the catch in an adjacent standard U. S. Weather Bureau gage. Solid line represents a perfect relationship.

ible for the large monthly quantities that usually fall in the wet mountain regions of Oahu.

Every attempt was made to locate the gages in accordance with the U. S. Weather Bureau recommendations. However, because of the very dense vegetation and the rough terrain, it was not always possible to follow those suggestions completely.

*Climate of Oahu.* For a relatively small oceanic island, Oahu supports an unusually large population and maintains a highly organized agricultural economy; consequently, its climate has been the subject of many investigations, particularly with reference to rainfall. The most comprehensive attempt to describe the pattern of distribution of rainfall on Oahu was made in a monograph [Leopold and others, 1951] devoted to the rainfall of the Hawaiian Islands. More recently an intensive investigation was made of the mechanism of rain formation in a part of the island of Hawaii under the Project Shower [Mordy and others, 1957]. Many of the results of that work are applicable also to the mountainous region of Oahu.

The monograph mentioned above [Leopold

and others, 1951] clearly describes the climatology of the Koolau Range, and the reader is referred to it for details. Briefly, the rainfall of the Hawaiian Islands is normally related to either trade-wind or cyclonic circulation. The wind circulation, occurring throughout most of each year, is the more general case, and quantitatively it is the most important to rainfall in the wet mountains. Typical trade winds blow from the east-northeast, approximately normal to the trend of the summit of the Koolau Range. The warm, moist air of the trade winds, cooled rapidly as it is lifted up the steep mountain, produces orographic rainfall. Where the mountains are less than about 5000 feet high, as is true of the Koolau Range, the air mass passes over the summit and the rainfall reaches a maximum some distance leeward of the summit. On the other hand, cyclonic circulation is characterized by a winter maximum and a summer minimum and its associated air mass generally flows from a southerly direction. This circulation, locally known as 'Kona' weather, typically results in humid, sultry weather, often accompanied by storms with heavy rainfall. Rainfall from these storms is rather uniformly distributed over wide areas and accounts for most of the rain that falls in the leeward coastal regions.

The highly variable areal distribution of rainfall on Oahu has long been recognized. Even a casual observer watching the march of an orographic shower over the Koolau Range can notice that the intensity of the shower apparently decreases with increasing distance from the summit. Stidd and Leopold [1951] pointed out that the annual isohyets follow a geometric progression, increasing toward the higher elevations on Oahu. This geometric progression is controlled by the orographic-rainfall component, since cyclonic rains are relatively uniformly distributed. Baer [1956] attempted to explain the phenomenon of the orographic-rainfall pattern on the basis of meteorological parameters and derived an equation showing its logarithmic distribution. During the present study it became obvious that the pattern of rainfall distribution in the central leeward Koolaus could be closely correlated with distance from the summit. This relationship had also been suggested by Vood [1928].

TABLE 2. Total 1957 Rainfall on Approximate Monthly Basis (in inches)

| Interval           |                     |                    |                    |                    |                   |                    |                     |                    |                    |                    |                            | Total  |
|--------------------|---------------------|--------------------|--------------------|--------------------|-------------------|--------------------|---------------------|--------------------|--------------------|--------------------|----------------------------|--------|
| Jan. 2-<br>Jan. 28 | Jan. 28-<br>Feb. 26 | Feb. 26-<br>Apr. 3 | Apr. 3-<br>Apr. 30 | Apr. 30-<br>May 28 | May 28-<br>Jul. 1 | Jul. 1-<br>Jul. 30 | Jul. 30-<br>Sept. 3 | Sept. 3-<br>Oct. 1 | Oct. 1-<br>Oct. 31 | Oct. 31-<br>Dec. 3 | Dec. 3-<br>Jan. 2,<br>1958 |        |
| 18.38              | 4.69                | 1.70*              | 4.88               | 2.56               | 1.81              | 3.16               | 8.00                | 0.75               | 0.85               | 8.80               | 5.25                       | 60.83  |
| 19.00*             | 6.69                | 1.94               | 5.94               | 2.06               | 2.50              | 4.44               | 10.91               | 0.90               | 1.10               | 10.31              | 8.60                       | 74.39  |
| 19.13              | 7.69                | 2.45*              | 7.20               | 2.25               | 3.15*             | 5.31               | 12.00               | 1.56               | 2.13               | 11.13              | 8.70                       | 82.70  |
| 19.06              | 7.50                | 3.25               | 8.22               | 2.88               | 3.31              | 6.97               | 13.50*              | 1.75               | 2.38               | 11.31              | 9.00                       | 89.13  |
| 18.31              | 9.63                | 3.00               | 8.00               | 3.75               | 4.50              | 7.94               | 15.56               | 2.50               | 2.69               | 12.13              | 10.00                      | 98.01  |
| 18.25              | 11.25               | 3.25               | 9.69               | 4.80*              | 5.00              | 8.69               | 17.25               | 3.06               | 2.75               | 15.25              | 12.78                      | 112.02 |
| 20.56              | 11.75               | 3.50               | 12.00              | 6.25*              | 6.35              | 11.75              | 19.50               | 4.38               | 4.50               | 20.94              | 14.81                      | 136.29 |
| 18.50              | 12.63               | 4.00               | 13.63              | 8.50               | 7.00              | 14.50              | 22.25               | 5.44               | 5.56               | 24.13              | 16.16                      | 152.30 |
| 20.19              | 12.63               | 3.69               | 16.63              | 8.78               | 7.35              | 15.00              | 22.88               | 6.13               | 7.19               | 29.81              | 16.72                      | 167.00 |
| 21.41              | 12.81               | 3.88               | 20.50              | 8.72               | 7.63              | 16.38              | 23.63               | 7.00               | 9.31               | 37.50              | 17.38                      | 186.15 |
| 18.69              | 10.44               | 3.44               | 21.16              | 6.63               | 6.16              | 12.94              | 19.28               | 6.25               | 9.56               | 31.75              | 14.31                      | 160.61 |
| 18.63              | 11.00               | 3.81               | 18.94              | 5.75               | 6.38              | 13.47              | 21.75               | 6.25               | 10.00              | 36.06              | 14.88                      | 166.92 |
| 20.06              | 9.37                | 2.81               | 18.10              | 4.50               | 5.13              | 10.38              | 16.25               | 5.25               | 8.50               | 28.25              | 11.81                      | 140.41 |
| 19.63              | 7.44                | 2.88               | 17.00*             | 3.88               | 4.31              | 9.34               | 14.44               | 4.47               | 7.56               | 25.25              | 10.69                      | 126.89 |
| 18.66              | 6.00*               | 1.94               | 5.46               | 2.40               | 1.99              | 2.95               | 9.00                | 0.75               | 0.80               | 10.28              | 8.03                       | 68.26  |
| 20.20              | 9.74                | 1.90               | 7.81               | 2.39               | 3.60              | 6.10               | 12.74               | 1.88               | 2.00               | 14.49              | 12.20                      | 95.05  |
| 20.41              | 10.65               | 3.27               | 9.15               | 2.64               | 4.88              | 8.44               | 15.25               | 2.75               | 2.63               | 15.45              | 14.05                      | 109.57 |

\* Value interpolated because of faulty reading.

Results of the present work are based on observations made at monthly intervals throughout 1957. Climatologically, 1957 was not a normal year for Oahu, but it is likely that the distribution of rainfall, although not the amount, followed a nearly normal pattern. The U. S. Weather Bureau's 1957 summary of climatological data for Hawaii [Blumenstock, 1958] notes that 10 months of the year were unusually dry and 2 months were unusually wet, and during the remaining 4 months precipitation was very close to the long-term average. In the leeward Waianae Range the 1957 rainfall was below normal. The long-term annual average in the vicinity of gage 1 should be about 75 inches, according to the U. S. Weather Bureau's isohyets. The 1957 rainfall amounted to about 61 inches (see Table 2), or approximately 80 per cent of normal. However, since the distribution pattern of orographic rainfall for both the very dry and the very wet months was similar, it is concluded that the distribution pattern for the year was essentially normal.

**Results and Discussion.** From the plot of rainfall as a function of distance from gage 8 (Fig. 2), it is evident that the increase in rainfall to the point of maximum on the leeward

side of the summit follows an equation of the type:

$$\log y = ax + b$$

It is also evident, although less perfectly so, that the increase in rainfall as a function of distance from the summit to the leeward maximum can be described by an equation of similar type.

The rainfall is greatest near gage 10, about 2500 feet leeward of the summit where the abrupt change in topographic slope occurs. From this point up to gage 14 on a small peak on the summit ridge, the rainfall decreases rapidly until, at the summit, it is only about two-thirds as great as at gage 10. The erratic nature of this sharp decrease is probably due to turbulent spillover as the orographic cloud passes over the summit. The turbulence presumably is the result of a sharp pressure change caused by the occurrence of a hydraulic jump in the air passing over the ridge (L. B. Leopold personal communication, 1958). Frequent strong, variable winds may also influence the catch conditions in this region.

The two curves in Figure 4 show the relationship between (1) the log of the ratio of the rainfall at each gage to the rainfall at gage 8 for a selected period of orographic rainfall and



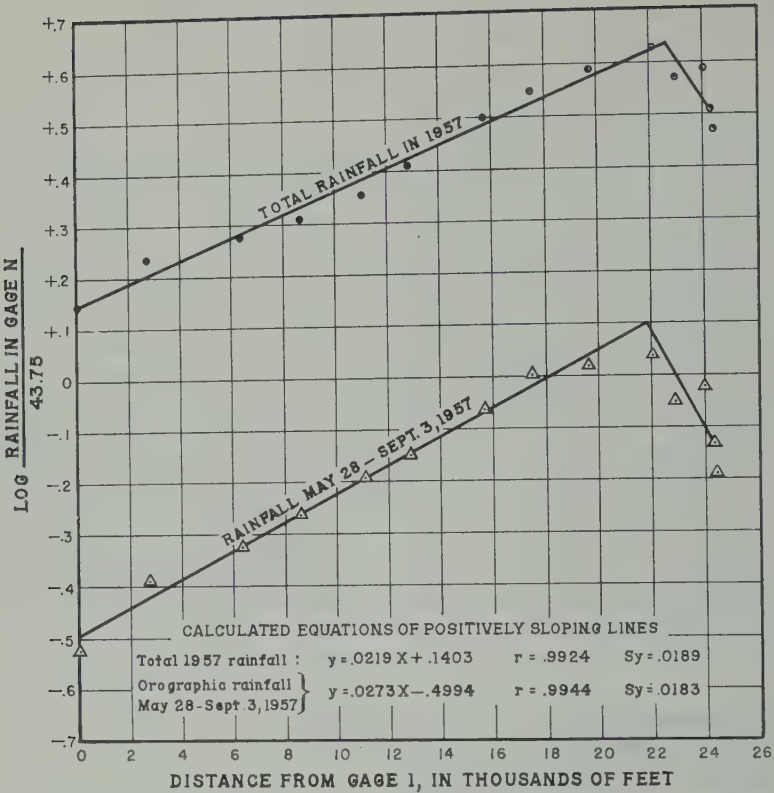
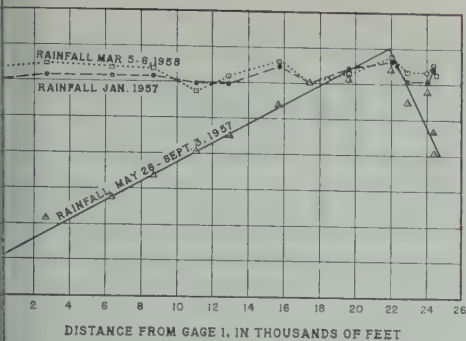


Fig. 4. Rainfall relationships on Kipapa Ridge plotted against distance from gage 1, during whole of 1957 and during a selected interval of orographic rainfall. Rainfall is expressed as the log the ratio of the catch at each gage to the catch at gage 8 during the orographic period, which amount to 43.75 inches.

(2) the distance from gage 1. One curve shows the relationship during the whole of 1957 and the other the relationship during the period of orographic rainfall in the summer of 1957. These curves illustrate convincingly that the annual pattern is dominated by the orographic rainfall component. *Stidd and Leopold* [1951] noted this effect in their general analysis of the geographic distribution of rainfall in the Hawaiian Islands. They stated that 'in the absence of orographic activity, this cyclonic activity would provide a uniformly thick blanket of rainfall at all gages. Orographic activity, on the other hand, works in such a fashion as to supply varying amounts of rainfall at different stations.' Gage 8 was chosen as the reference gage because of its intermediate location along Kipapa Ridge. The summer period of orographic rainfall consisted entirely of trade-wind days except for two nearly

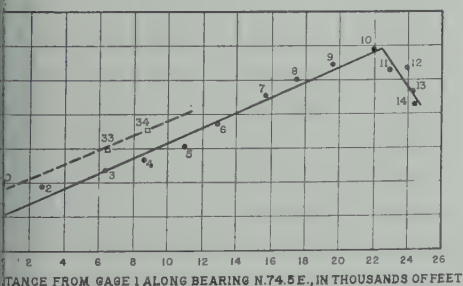
dry days in each of July and August. The difference between the slopes of the positively sloping sections of the two curves is attributed to aberrations from the blanketlike distribution of rainfall from cyclonic storms. The equations of the positively sloping sections of the curves were derived by the method of least squares, which the rainfall at gages 1 to 10 was used for the whole of 1957 and rainfall at gages 1 to 10 was used for the period of orographic rainfall. Gage 10 was not included in determining the regression line for the period of orographic rainfall because, for this particular interval, the point of maximum rainfall apparently lay between gage 9 and gage 10, whereas for the whole of 1957 it lay between gage 10 and gage 11. The correlation of the rainfall with distance from the summit is extremely high for both curves (Fig. 4) and the values of  $Sy$  (standard error



5. Comparison of cyclonic-storm rainfall and orographic rainfall on Kipapa Ridge. Rainfall is expressed as the log of the ratio of the catch at ridge gage to the catch at gage 8.

ate), converted to inches of rainfall, amount to only 4.5 per cent of the rainfall at any point on the line of regression for each period.

The comparative uniformity of distribution of rainfall during cyclonic storms is clearly indicated in Figure 5. Two curves in the figure represent the rainfall in two periods of cyclonic storms plotted as the log of the ratio of each gage to gage 8 against distance from gage 1. The curve from Figure 4 for the summer period of orographic rainfall is added to the figure for comparison. The quantities of rain that fell during the cyclonic storms and the period of orographic rainfall are listed in Table 3. The differences in catch among the gages during each cyclonic-storm period largely represent the effect of local conditions of exposure, but minor orographic effects associated with cyclonic air may also be significant.



6. Comparison of total rainfall on Kipapa Ridge with total rainfall in the bottom of Kipapa Valley in 1957. Rainfall is expressed as the log of the ratio of the catch at each gage to the catch at gage 8.

TABLE 3. Selected Trade-Wind and Cyclonic-Storm Rainfall Periods

| Gage     | Trade-Wind<br>Rainfall,<br>in inches<br>May 28-<br>Sept. 3, 1957 | Cyclonic-<br>Storm<br>Rainfall,<br>in inches<br>Jan. 2-28,<br>1957 | Cyclonic-<br>Storm<br>Rainfall,<br>in inches<br>March 5-6,<br>1958 |
|----------|--|--|--|
| Ridge 1  | 12.97  | 18.38  | 13.16  |
| 2        | 17.85  | 19.00  | 13.56  |
| 3        | 20.46  | 19.13  | 13.28  |
| 4        | 23.78  | 19.06  | 13.22  |
| 5        | 28.00  | 18.31  | 11.38  |
| 6        | 30.94  | 18.25  | 12.75  |
| 7        | 37.60  | 20.56  | 13.88  |
| 8        | 43.75  | 18.50  | 12.25  |
| 9        | 45.23  | 20.19  | 13.03  |
| 10       | 47.64  | 21.41  | 14.56  |
| 11       | 38.38  | 18.69  | 13.10  |
| 12       | 41.60  | 18.63  | 13.10  |
| 13       | 31.76  | 20.06  | 13.69  |
| 14       | 28.09  | 19.63  | 12.94  |
| Gulch 30 | 13.94  | 18.66  | 13.46  |
| 33       | 22.44  | 20.20  | 14.31  |
| 34       | 28.57  | 20.41  | 14.43  |

The gages on the floor of the gulch consistently recorded higher rainfall than those on the ridge at similar distances from gage 1. Gage 1 on the ridge makes a pair with gage 30 in the gulch; gages 3 and 33 make a second pair and gages 4 and 34 a third pair. Figure 6 shows that the rainfall in the gulch is significantly greater than that at comparable locations on the ridge. It shows also a geometric progression of isohyets similar to that demonstrated for the ridge. During 1957, gage 30 collected 7.43 inches of rain more than gage 1; gage 33 collected 12.35 inches more than gage 3; and gage 34 collected 20.44 inches more than gage 4 (see Table 2). The differences in rainfall between comparable locations in the gulch and on the ridge may be a result of subsidence of the trade-wind flow of air between the leeward ridges trending perpendicular to the summit ridge. This subsidence would result in a higher column of rain-laden air over the gulch than on the adjacent ridges. Wind effects might also tend to divert the rain from the ridges into the gulch. The differences in rainfall probably prevail up to the point of abrupt change in slope, about 2500 feet from the summit. From there to the summit the rainfall is

probably about the same at points in the gulch and at opposite points on the ridges.

*Conclusions.* On the wet leeward slopes of the Koolau Range of Oahu the rainfall distribution pattern can be correlated with distance from the north-northwestward-trending summit ridge. The dominant influence on the distribution pattern is orographic rainfall, since rainfall from cyclonic storms is relatively uniformly distributed. Rainfall in the bottoms of gulches is significantly higher than that on adjacent ridges, probably owing to subsidence of air flow and to wind effects. The results of the study are applicable to the small drainage basins in the west central leeward section of the Koolau Range. They may also be of significance in comparable locations elsewhere where a relatively constant flow of moist air passes over a mountain barrier. Further study would be required in order to determine whether the calculated equations are applicable to the drier leeward areas of Oahu, farther southwest from the crest of the range.

## REFERENCES

- Baer, Ledolph, Orographic rainfall distribution with application to Hawaii, *Trans. Am. Geophys. Union*, 37, 546-547, 1956.
- Blumenstock, D. I., Weather summary in *Climatological Data, Hawaii, Annual Summary*, 1953, 150, 1958.
- Leopold, L. B., and others, On the rainfall of Hawaii: A group of contributions, *Meteorol. Monographs*, 1(3), 55 pp., 1951.
- Mordy, W. A., and others, Project shower, and investigation on warm rainfall in Hawaii, *Tellus*, 9, 471-590, 1957.
- Stidd, C. K., and L. B. Leopold, The geographical distribution of average monthly rainfall, Hawaii, in 'On the rainfall of Hawaii: A group of contributions,' *Meteorol. Monographs*, 1(3), 242-251, 1951.
- Voorhees, J. F., A quantitative study of rainfall of the Island of Oahu, *Suppl. to Rept. of Honolulu Sewer and Water Comm., Appendix A*, 1927, 1928.

(Manuscript received May 6, 1960; revised June 22, 1960.)

## Evaporation Losses from Small-Orifice Rain Gages

HAROLD E. GILL

*U. S. Geological Survey  
Trenton 7, New Jersey*

**Abstract.** Evaporation losses from two types of small-orifice rain gages, a  $2.3 \times 2.5$ -inch wedge-shaped gage and a 1.15-inch gage, limit their accuracy when it is impossible or impractical to read them shortly after precipitation has ceased. Daily evaporation losses for the wedge-shaped gage ranged from 0.05 to 0.30 inch, with the maximum loss of 0.30 inch consistently recorded in the 4- to 6-inch precipitation range during the summer months. The 1.15-inch gage showed a maximum daily evaporation loss of 0.04 inch in the 4- to 6-inch precipitation range, whereas for rainfalls of less than 2 inches the daily evaporation losses were almost negligible. Although most meteorologists would be aware of the shortcomings of the wedge-shaped gage, the data presented might serve to alert workers in allied fields to errors caused by using them as storage gages.

**Introduction.** In the past few years, several small-orifice rain gages have appeared on the market. They are much cheaper than the standard U. S. Weather Bureau 8-inch stick gage and easy to install and maintain.

For a hydrologic investigation in Cape May County, N. J., two types of small-orifice gages were installed to supplement the base network precipitation stations in the county. During the period, excessive evaporation losses were noted at ten stations where the small-orifice gages were installed, and the present study was undertaken to evaluate the effectiveness of small-orifice gages in supplying reliable data for hydrologic studies.

During the spring and summer of 1958, the evaporation losses from the two types of small-orifice gages were recorded at a test site near Trenton, N. J. This investigation was undertaken to determine the effect of gage size and shape on evaporation losses. Several investigators have previously studied the effects of gage size on precipitation catch. Among these are Conover and Nastos [1951], who tested 3- and 1.15-inch small-orifice gages, and Huff [1951], who compared the 3-inch,  $2.5 \times 2.3$ -inch wedge-shaped, 1.15-inch tube, and the 0.8-inch glass gages with the U. S. Weather Bureau 8-inch stick gage. These studies indicated that the evaporation losses of the small-orifice rain gages compared favorably with that of the standard gage.

**Description of gages.** The plastic  $2.5 \times 2.3$ -

inch wedge-shaped gage has a rectangular catch area tapering to a wedge-shaped well. The scale is imprinted on the side and is graduated at intervals of 0.01 inch to a depth of 0.2 inch, then at 0.02-inch intervals to 1-inch and 0.05-inch intervals above. This change in graduation is required by the shape of the gage. The gage, which has a capacity of 6 inches, is 13.5 inches long and comes equipped with a metal bracket for mounting on a fence or post.

The 1.15-inch tube gage is a plastic tube of 0.67-inch inside diameter with an aluminum funnel inserted in the tube opening. The gage is designed to magnify the catch 3 times. The tube is mounted in an aluminum frame suitable for mounting on a fence or post. A scale graduated at 0.10-inch intervals is etched on the aluminum frame. The gage has a capacity of 6 inches and is approximately 19 inches long.

**Field installation.** Two wedge-shaped gages and a 1.15-inch gage were operated simultaneously with a standard 8-inch stick gage from May to October 1958. Because the plastic gages cannot be used effectively during freezing weather, the evaporation studies were not begun until all threats of freezing temperatures had passed.

The rain gages were installed on the side of a small knoll in a residential area in Fairless Hills, Pa. The distance between the gages and any obstruction was at least  $2\frac{1}{2}$  times the height of the obstruction. The buildings offered some pro-



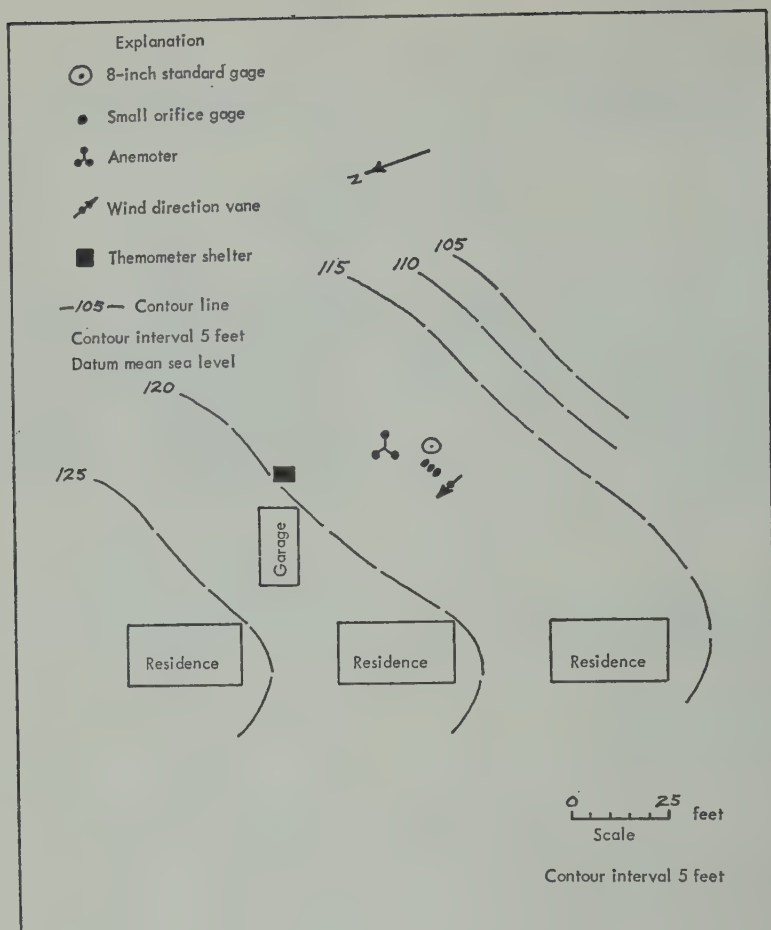
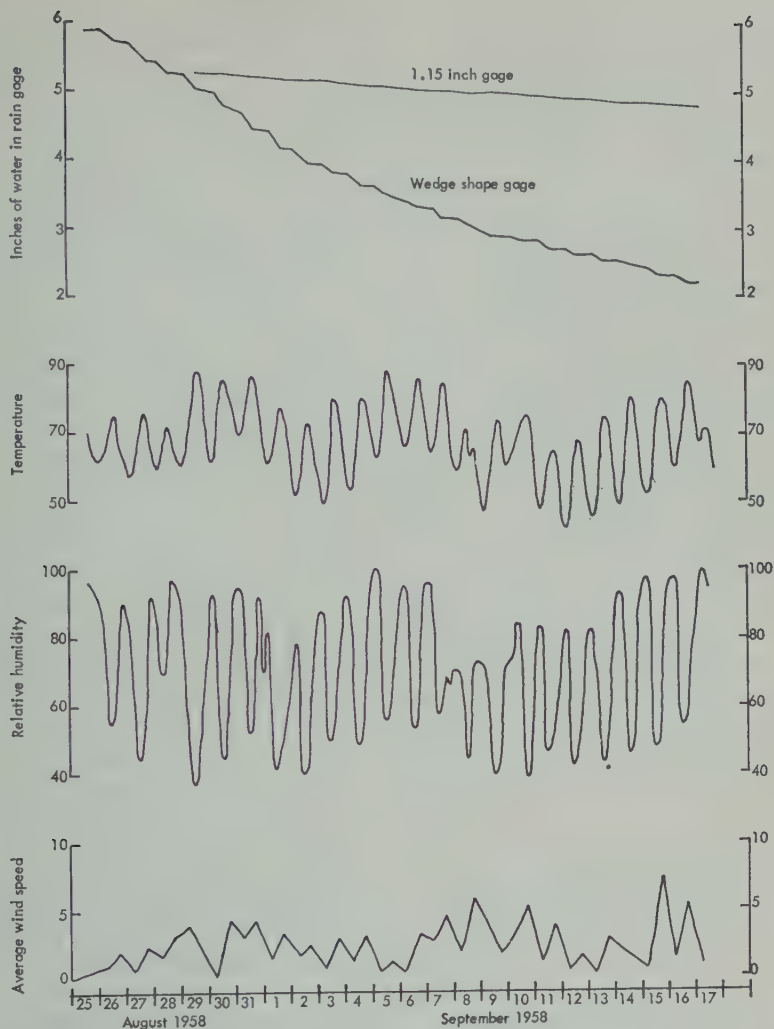


Fig. 1. Map of test plot in Fairless Hills, Pa., showing location of instruments.

tection from the wind, so that the wind velocity near the orifices of the rain gages rarely exceeded 10 miles per hour. This location was chosen (1) to minimize one of the major factors controlling evaporation, wind velocity, and (2) because of the necessity of having an installation close enough to the observer so that frequent observations of weather conditions could be made.

The three small-orifice gages were installed at a distance of 3 feet from the standard gage and at the same height of approximately 36 inches (Fig. 1). An anemometer, equipped with a 3-cup totalizing odometer, was installed at a distance of 12 feet from the standard gage, about 14 feet from the small-orifice gages, and at the

same height as the orifices of the gages. A thermograph and barograph and a maximum and minimum thermometer were operated in a standard shelter at a distance of 35 feet from the gages. Readings were made twice daily at 7:30 A.M. and 6:00 P.M. The data recorded included: maximum and minimum temperature at time of observation, barometric pressure, average wind speed for period of observation, wind direction, dew point, relative humidity, and evaporation loss from each gage. The 1.15-inch gage, capacity 6 inches, one of the wedge-shaped gages, capacity 1.15 inches, were filled to near capacity at the beginning of the evaporation studies. The second wedge-shaped gage was filled to a depth of



2. Evaporation losses of small-orifice rain gages compared with temperature, humidity, and wind speed during the period August 25 to September 17, 1958.

to expose a smaller surface area. Because purpose of the investigations was to determine the quantity of water lost from the small-orifice rain gages by evaporation with time, the gages were artificially filled to simulate catches of storms of various magnitudes.

**Analysis of data.** The data from each of the small-orifice gages were plotted separately for the longest dry periods of record. A composite graph (Figure 2) shows the daily amount of water in the 1.15-inch and wedge-shaped rain gages during the period August 25 to September

17, 1958. The wedge-shaped gage with its relatively large surface area had the most evaporation loss. As the water level fell in the wedge-shaped well, the rate of evaporation diminished. However, even when the water level in the wedge-shaped gage was below the 1-inch level, (Fig. 3), the rate of evaporation loss still was approximately twice the maximum loss from the 1.15-inch gage in the 5- to 6-inch precipitation range.

Graphs of the hourly temperature, humidity, and average wind speed were plotted for corre-

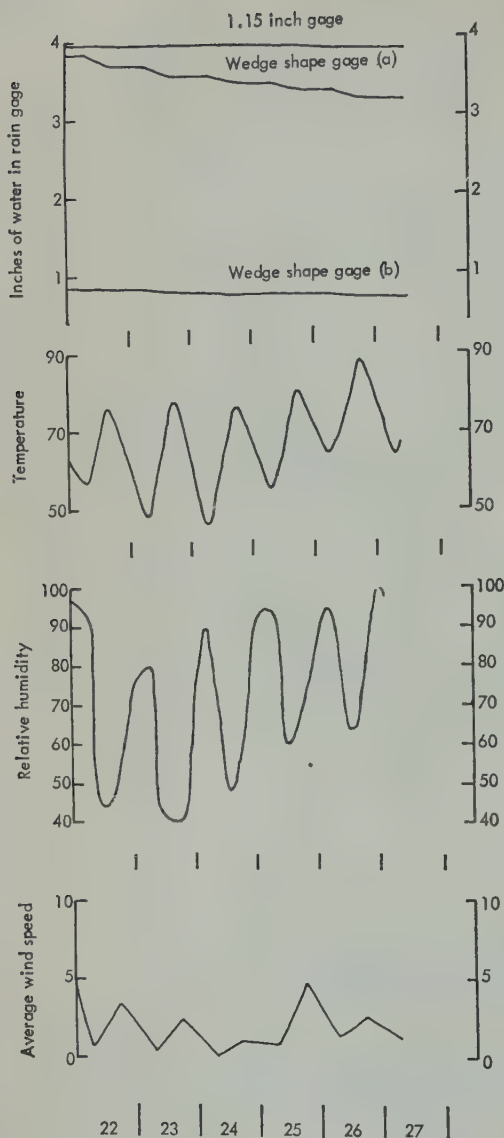


Fig. 3. Evaporation losses of small-orifice rain gages compared with temperature, humidity, and wind speed during the period September 22 through September 27, 1958.

lation with the evaporation losses from the rain gages for the period August 25 to September 17 (Fig. 2) and September 22 to 27 (Fig. 3) to show the magnitude of evaporation loss for rainfalls in the 1- to 6-inch range. Comparison of the relative humidity obtained with a standard sling psychrometer at the test plot with the

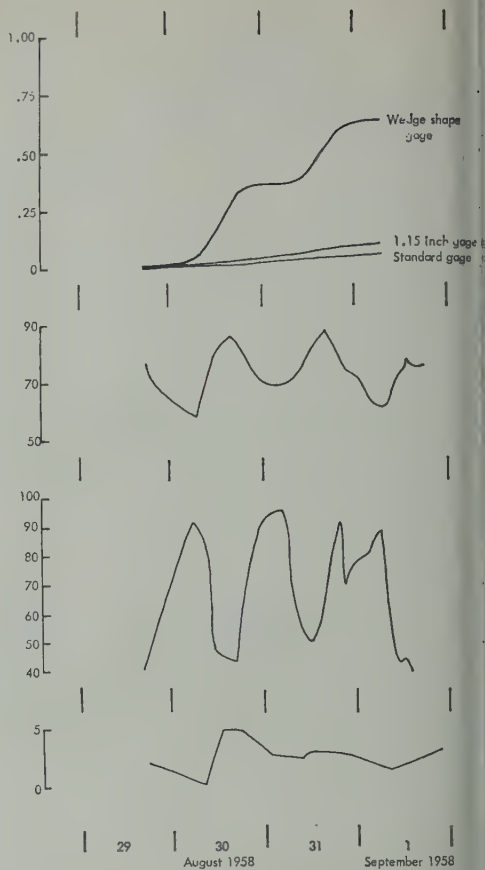


Fig. 4. Cumulative hourly evaporation loss of small-orifice rain gages compared with temperature, humidity, and wind speed.

U. S. Weather Bureau hygrograph at Trenton, N. J., for comparative periods during August and September 1958, showed a close correlation. For this reason the continuous hygrograph record maintained by the Weather Bureau at Trenton was used to supplement the test data. The average wind speed was obtained for the period of observation and plotted as a point on the graph representing the preceding 12 hours.

Temperature, relative humidity, and wind speed were the apparent controlling factors in evaporation losses from the several rain gages. Temperature and relative humidity were the main factors, as the area was afforded some protection from the wind. The average 12-hour wind speed at a height of 36 inches generally varied from 0 to 10 miles per hour.

imum daily water loss of 0.30 inch was recorded consistently for the wedge-shaped gages for precipitation catches of between 4 and 6 inches during the summer of 1958. In the same time period (Figs. 2 and 3) evaporation losses were between 0.10 and 0.20 inch in the 3-inch precipitation range, whereas the daily loss was between 0.10 and 0.15 inch for catches between 2 and 3 inches. For precipitation catches of 1 inch or less, the maximum daily evaporation loss recorded was 0.05 inch.

imum daily water loss of 0.04 inch was recorded for the 1.15-inch gage for precipitation catches of between 4 and 6 inches but were negligible for precipitation catches of 2 inches or less. In an open area with similar temperature and relative-humidity conditions as at the test site where the wind could reach much higher velocities than the 12-hour average of 10 miles per hour, larger evaporation losses might be expected.

From August 29 to September 1, 1958, more frequent measurements were made to determine at which portion of the day the major evaporation losses occurred. Figure 4 shows a cumulative plot of the evaporation from the 1.15-inch and wedge-shaped gages. Both the 1.15-inch and wedge-shaped gages were filled to show the effect of precipitation. The wedge-shaped gage recorded the greatest evaporation loss between 6:00 A.M. and 7:00 P.M., and the 1.15-inch gage registered its maximum evaporation loss in the late afternoon. Small losses were recorded during the evening hours during this period from both small-orifice gages, especially from August 31 through September 1, when the precipitation remained low and the wind velocity was above normal.

*Summary and conclusions.* Results of the investigation indicate that there are serious limitations to the use of the wedge-shaped rain gage where it is impractical or impossible to read the instrument shortly after rainfall. Although the results of previous studies [Huff, 1955] have shown that the catch of the wedge-shaped gage when compared with the standard U. S. Weather Bureau gage is somewhat more accurate than that of the 1.15-inch gage, the higher evaporation loss from the wedge-shaped gage limits its use in hydrologic studies. Conover and Nastos [1951] expressed the desirability of using a small-orifice 1.15-inch gage when only rain is to be measured. They found that the small size of the gage disturbs the wind less, and liquid precipitation catches recorded are greater than those measured by the standard 8-inch U. S. Weather Bureau gage (3% at 30 mph wind speed). It was their opinion that the small-orifice gage recorded a truer picture of the precipitation in their area of investigation. This fact and the comparatively low evaporation loss make the 1.15-inch gage a practical and inexpensive instrument for supplementing U. S. Weather Bureau precipitation data for hydrologic studies.

#### REFERENCES

- Conover, J. H., and T. J. Nastos, Test of Stewart and Victor precipitation gages, *Blue Hill Rept., Contract CWB-8120*, Harvard University, Blue Hill Meteorological Observatory, Milton 86, Mass.
- Huff, F. A., Comparison between standard and small orifice raingages, *Trans., Am. Geophys. Union*, 36, 689-694, 1955.

(Manuscript received March 28, 1960; revised May 23, 1960.)





# Percolation Measurements Based on Heat Flow Through Soil with Special Reference to Paddy Fields

SEITARO SUZUKI

*The Training School for Meteorological Observers  
Kashiwa City, Chiba Prefecture, Japan*

**Abstract.** A new method is presented for the determination of the *in situ* percolation rate of soil based on the influence percolating water has on the heat flow through soil.

**Introduction.** The shortage of irrigation water in paddy fields in summer is of deep concern to rice-producing countries. In general, a rice-cultivated field loses water in three ways: evaporation, transpiration, and percolation. The separate measurement of each loss is useful in obtaining information on factors involved and reducing losses to a minimum. With formulas of V. Thornthwaite and B. Holzman the total amount of transpired and evaporated water is estimated, but the two effects cannot be compared separately. Penman [1958] mentions that Rothamsted Experimental Station has equipment which gives a continuous record of transpiration. A paper on 'Measurements of Evaporation and Transpiration by Rice in Paddy Fields' is being published separately by the author. This paper is concerned with a new method for measuring the percolation rate.

**Measurement of Percolation.** The percolation rate of water in paddy field has been measured by the 'percolation meter.' This consists of a long tube of fairly large diameter, with a cap to prevent evaporation. It is injected into the soil to its full length and is filled with water to nearly the same level as the ambient irrigation water. (It has come to the attention of the author that American investigators have been using infiltrometers of the single- or double-ring type to measure percolation rate [Schiff, 1953; and Luthin, 1957; Haise, Donnan, Phelan, et al., and Shockley, 1956]. This apparently has advantages over the Japanese percolator). The rate of water-level drop in the tube is assumed to be the same as the percolation rate of water in the paddy field around the tube. This assumption is true when the soil has no transpir-

ing rice plants outside the tube. Less infiltration occurs outside the percolation tube due to use of water by embedded plant roots. Thus it is less than that measured by means of the percolation meter. The sharp edge of the meter cuts plant roots loose and prevents root activity inside. K. Tanabe has shown that percolation losses are reduced as the vegetative cover becomes thicker and transpiration increases. It is difficult to determine the water percolating through the root zone without lysimeters in the plot. However, it is believed that the following device will offer a more practical means of measuring the amount of water percolating in a paddy field *in situ*. The device is not expensive and it does not disturb the soil. It involves the principle that the conduction of heat in the soil of a paddy field increases (or decreases) not only with the conductivities of the soil constituents but also with the percolation rate of warm (or cold) irrigating water. A mathematical expression of this relationship is simply derived as follows.

Suppose that heat as well as water flows only vertically upward or downward (along the  $z$  axis). Then the following thermal equation is applicable for the determination of temperature distribution of soil, provided the percolation rate of water is sufficiently small:

$$\frac{\partial \theta}{\partial t} = \frac{k}{c\rho} \frac{\partial^2 \theta}{\partial z^2} - \frac{vc_0\rho_0}{c\rho} \frac{d\theta}{dz} \quad (1)$$

where  $\theta$  represents the temperature of the soil at the depth  $z$ ,  $k$  represents the heat conductivity,  $c$  the specific heat,  $\rho$  the apparent density of the saturated soil, and  $v$  the vertical percolation rate of water through a unit area per unit time, the density and specific heat of water  $\rho_0$ ,  $C_0$  be-

ing assumed to be unity. The last term on the right side of (1) represents the only modification to the classic equation of the thermal soil conduction. If the surface temperature of the soil is a simple harmonic variation, the solution is

$$\theta = \theta_0 + Ae^{-az} \sin(2\pi t/T - bz) \quad (2)$$

where

$$a = (\pi c \rho / Tk)^{1/2} - v/2k \quad (3)$$

$$b = (\pi c \rho / Tk)^{1/2} \quad (4)$$

$\theta_0$  = the mean temperature.

$A$  = the amplitude of temperature wave at the soil surface.

Both of the above formulas are valid precisely to the first order of  $v/2k$ . It may be helpful to point out that  $a$  (logarithmic decrement of temperature) is obtainable in the usual way by determining both maximum and minimum temperatures  $\theta_1^+$  and  $\theta_1^-$ ,  $\theta_2^+$  and  $\theta_2^-$  at two depths  $z_1$  and  $z_2$ , respectively:

$$a = [\ln(\theta_1^+ - \theta_1^-) - \ln(\theta_2^+ - \theta_2^-)] / (z_2 - z_1) \quad (5)$$

Another term,  $b$ , can be determined if the phase difference of the temperature wave between two depths is known (the difference of occurrence-times of the maximum or minimum temperature  $t_1$  and  $t_2$  at  $z_1$  and  $z_2$  respectively).

$$b = 2\pi(t_2 - t_1) / T(z_2 - z_1) \quad (6)$$

The combination of (3) and (4) gives the percolation rate sought:

$$v = 2k(b - a) \quad (7)$$

where the soil heat conductivity  $k$  is the only unknown quantity to be determined. This quantity may be found either directly by a physical experiment or indirectly by arithmetically summing up all heat conductivities of the soil constituents by references to physical tables. However, the writer has attempted to find  $C\rho$  (volumetric specific heat of the sampled soil)

by a calorimetric method. The soil conductivity can be computed from (4) with the observed value of  $C\rho$ , with  $b$  obtained from (6), and with a period of 1 day for  $T$ .

To test (4) *K. Yabuki's* [1951] study of the numerical relation between the distorted temperature-distribution and the rate of percolation of a leaking plot was used. The computed values of percolation rate are compared with the observed values in Table 1 together with logarithmic decrements.

Kanto sandy loam soil, rich in volcanic ash, was used in these experiments. Its calorimetrically measured value of  $C\rho$  was found to be 0.119 cal/cm<sup>3</sup>; thermometric observation gave  $b = 0.0033$  cal, cm/cm<sup>2</sup>. Therefore

$$v = (7.829 - 65.79a) \times 10^{-4} \text{ cm}^3/\text{cm}^2 \text{ s}$$

This equation was used to compute the theoretical values of  $v$  given in Table 1.

From Table 1, it may be seen that, on the basis of the percolation experiment of Yabuki, the theoretical percolation rate approaches the actual one as the percolation rate increases. For example, the measured percolation rate of  $1.0 \times 10^{-5}$  cm<sup>3</sup>/cm<sup>2</sup> s is almost identical with the theoretical. This apparatus seems to provide a satisfactory method for making percolation measurements for the more permeable soils. However, the preceding experiments were not sufficiently accurate to justify a strict comparison between theory and experiment, and more thorough experimental research on this matter is desired.

This type of thermal percolation gage and

TABLE 1. Observed and Theoretical Percolation Rates and Logarithmic Decrements of Temperature

| $a$ (logarithmic decrement) | Observed Percolation Rate, cm <sup>3</sup> /cm <sup>2</sup> s | Theoretical Percolation Rate, cm <sup>3</sup> /cm <sup>2</sup> s |
|-----------------------------|---|--|
| 0.116                       | $0.0 \times 10^{-5}$  | $2.0 \times 10^{-5}$   |
| 0.110                       | 9.4   | 5.9  |
| 0.109                       | 5.6   | 6.6  |
| 0.099                       | 13.3  | 13.2   |
| 0.091                       | 16.6  | 18.4   |

dy explained tubular percolation meter will, properly combined, give two kinds of water s in a paddy-field, provided that the total unt of water supplied is simultaneously sured by some means such as a hydraulic attached to an irrigator.

us if  $F$ ,  $P$ , and  $D$  have the following nings,

= water loss due to percolation through paddy field per unit time (measured by a thermal method)

= water loss through paddy field soil per unit time (measured by an inserted tube)

= water supplied to field per unit time,

= transpiration =  $P - F$

= evaporation =  $D - P$ .

## REFERENCES

- Burgy, R. H., and J. N. Luthin, Discussion of 'A Test of the Single- and Double-Ring Types of Infiltrimeters,' *Trans. Am. Geophys. Union*, 38, 260-261, 1957.
- Haise, H. R., W. W. Donnan, J. T. Phelan, L. F. Lawhon, and D. G. Shockley. The use of cylinder infiltrimeters to determine the intake characteristics of irrigated soils, *Publ. of Agr. Research Serv. and Soil Conserv. Serv.*, U. S. Dept. Agr., pp. 1-10, May 1956.
- Penman, H. L., Discussion on water losses from various land surfaces, *Quart. J. Roy. Meteorol. Soc.*, 84, 190, 1958.
- Schiff, L., The effect of surface head on infiltration rates based on the performance of ring infiltrimeters and ponds, *Trans. Am. Geophys. Union*, 34, 257-266, 1953.
- Yabuki, K., On the temperature of the ground and water in the percolating paddy field, *J. Agr. Meteorol., Tokyo*, 7, 13-14, 1951.

(Manuscript received March 4, 1960; revised June 26, 1960.)





## Diffusion Effects in Miscible Displacement Occurring in Saturated and Unsaturated Porous Materials

J. W. BIGGAR AND D. R. NIELSEN

*Department of Irrigation  
University of California  
Davis, California*

**Abstract.** Miscible displacement has been studied in several porous materials, saturated and unsaturated, at different average flow velocities. The tracer appeared at the end of the soil column well in advance of what would have been expected had no mixing occurred at the boundary of the tracer and tracer-free water. Physical differences between porous materials were manifested by changes in shape and position of breakthrough curves due to ionic diffusion. One of the more important physical features was the magnitude of the volume of water not readily displaced at saturation and its increase when the soil was desaturated. Because the total flux of water moving through field soils is generally small, the role of hydrodynamic dispersion and diffusion in transporting dissolved solutes must be included in the theory of most soil-water processes.

**Introduction.** The spreading of a tracer in a porous medium during miscible displacement was first observed by Slichter [1905], who studied groundwater flow in aquifers. Slichter recognized that the velocity distribution in the pores led to a spreading of the tracer. Because the rate of flow was high, Slichter concluded that molecular diffusion was negligible in the transport of the tracer from the point of injection. Yuhara [1954] gave a mathematical interpretation of dispersion resulting from variations of local velocity vectors from the average velocity vector in a porous medium.

Later, Kaufman, and Todd [1956] and Scheidegger [1954, 1958] proposed statistical models to account for the spreading of the front during miscible displacement of the fluid in media. Rifai [1958] stated that transverse molecular diffusion introduced variation in the dispersion of the front. In Scheidegger's statistical model, based on an assumption of complete disorder of the porous material, only velocity distribution was considered to cause a spreading of the moving front. Molecular diffusion as a result of concentration gradients was ignored.

Using the Scheidegger model, Day [1956] explained the experimentally measured spreading of a salt-water boundary moving through sand and exchange resins. The theory accurately predicted the concentration distribution

within the columns; however the residence time was found to be in error.

Working with three soils, Oakley sand, Yolo sandy loam, and Hesperia sandy loam, Kaufman and Orlob [1956] investigated the use of several tracers as a means of tracing water movement. Chloride ion appeared to be the best tracer, with tritium, fluorescein, dextrose,  $I^{131}$ , and  $Ca^{45}$  showing varying amounts of exchange or adsorption. Since their primary objective was to determine which tracers were most suitable for tracing water movement, the interactions of the tracer and porous material were not fully investigated.

Another statistical model based on flow velocity distribution has been proposed by de Josselin de Jong [1958]. The model is a schematic pore canal system and the movement of the liquid through the canals is accounted for in detail. The probability distribution of the particle is continuous with respect to the direction of the path rather than a discontinuous distribution found by other workers. The theory combines both transverse and longitudinal dispersion into one formula, since it has been found that longitudinal dispersion is 6 to 8 times the magnitude of transverse dispersion. Calculations based on experimental data indicated that ionic diffusion was negligible in determining the shape of the tracer front.

Although the effect of molecular or ionic dif-

fusion in determining the spreading of a tracer front moving through porous media has been treated in some theories [Taylor, 1954; Rifai, Kaufman, and Todd, 1956], the assumption is made in most experimental work reported that diffusion effects are small or negligible since the average flow velocity is several magnitudes greater than the diffusion coefficients of the tracer in pure solutions. Von Rosenberg [1956], on the other hand, found qualitatively that convection and diffusion mechanisms are operating in controlling the shape of the front during miscible displacement. Using two tracers of different molecular diffusivity values, Handy [1959] showed by experiments at several flow velocities the effects of diffusion on the dispersion at slower flow rates.

Beran [1955] and Saffman [1960], using capillary models, were concerned with dispersion for cases in which molecular diffusion and microscopic mixing were both important. Both gave conditions in which one mechanism became dominant over the other. However, little experimental testing of the theories was presented. Although it is said that the theories may have importance in describing a model for porous media, experimental observations do not substantiate the validity of the theory when applied to granular beds.

In ground-water aquifers, soil materials, and other porous materials there exists an inherently more complex pore structure than in the washed single-grained sand and consolidated sandstones commonly used in the laboratory by previous investigators. In the former structures the pore size distribution is greater and the materials not only possess irregular shaped pores but also may contain more or less 'dead end' pores. These pores act as sinks for tracer diffusion. Here a more complete understanding of the porous material manifests itself through the breakthrough curves obtained at extremely small flow velocities, as will be shown later.

In addition to studies involving slow flow velocities, the contribution of ionic or molecular diffusion is further demonstrated when the water in the porous material is under negative pressure and a continuous air phase exists. Previous experimental attempts to maintain a negative water pressure under stationary flow conditions with a continuous but stagnant air phase have

proved unsuccessful. Orlob and Radhakrishna [1958], using positive pressure and a discontinuous air phase, investigated the effects of trapped air on the dispersion characteristics of sand. The hydraulic-medium dispersion constant  $D_m$  was drastically changed for different volumes of entrapped air. However, with the use of positive pressures all the curves were normal sigmoid-shaped curves similar to those found under saturated conditions.

The purpose of the present work is to provide basic experimental information regarding the combined role of molecular diffusion and hydrodynamic dispersion involved in miscible displacement. Intrinsic differences among porous materials will be manifested through the shapes and positions of the breakthrough curves obtained from saturated and unsaturated porous media.

*Experimental procedure.* An apparatus was specially designed so that stationary flow conditions and water content of the media could be maintained when tracer-free water was replaced by water containing a dissolved tracer. It was desired that no mixing of the two waters occur initially at the boundary between them. At the same time the contribution of flow associated with the boundary between soil and soil container had to be eliminated. Also, samples of effluent to be analyzed for tracer concentration had to be collected without disturbing stationary flow conditions. A cross-sectional sketch of the apparatus is given in Figure 1. Samples of air-dry soil screened through a U. S. Bureau of Standards No. 10 sieve (2mm) were uniformly packed in the 30-cm-long lucite tube. The water content and the volume flux density were

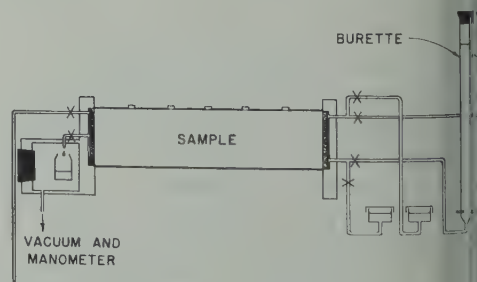
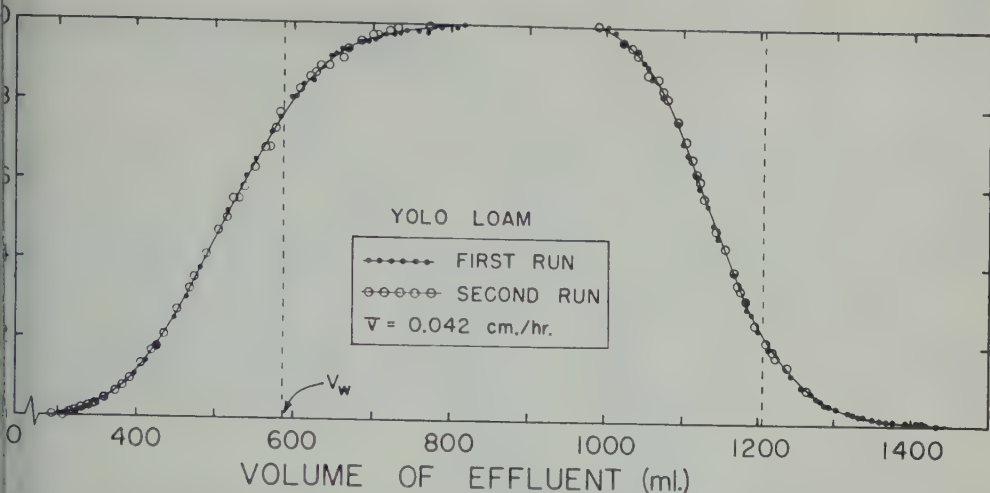


Fig. 1. Schematic cross section of apparatus used for collecting samples of effluent from columns maintained at constant water content and average flow velocity.



2. Relative chloride concentration  $C/C_0$  of effluent from Yolo loam. A 620-ml slug of chloride water was preceded and followed by tracer-free water in a soil column having a volumetric water  $V_w$  of 588 ml.

lled by using hanging water columns and we air pressure imposed on the fritted-head porous plates shown on each end of horizontal column. The use of fritted-glass-plates offers the advantage over other materials of having a negligible exchange capacity, large capillary conductivity, and an unusually narrow pore size distribution.

water in the small reservoir behind the plate could be rapidly changed from water containing a tracer to tracer-free water or the reverse by the use of leveling bottles. Burettes graduated to 0.2 ml were connected in the system to measure the volumes of water which entered and left the column. The X's in the figure designate necessary on-off valves. The vacuum arrangement allowed samples of effluent to be collected as close as possible to the end of the column, thereby eliminating further dispersion of solutes outside the sample. During the time the samples were not being collected, water was maintained by means of negative pressure obtained from hanging water columns at the same pressure as the vacuum arrangement. The necessary valves for this operation are in series with the effluent side of the soil column.

The negative water pressure over the entire length of column was of almost the same value, although small differences in pressure were allowed for flow. At the greatest average nega-

tive pressures used, the difference in pressures at the ends of the column never exceeded an equivalent of 5 cm of water. All studies were made under controlled temperature conditions. The volume of water in each sample was obtained by oven drying at 110°C.

Three soils, Oakley sand, Yolo loam, and Columbia silt loam, and two sizes of uniform glass beads, 200  $\mu$  and 390  $\mu$  were used in the investigation. The tracer-free water was boiled water with 0.01 N  $\text{CaSO}_4$ . The tracer water was boiled water of 0.1 N  $\text{CaCl}_2$  concentration. Analysis of the chloride ion was by titration with  $\text{AgNO}_3$ . Problems involving deflocculation and microbial activity were negligible because a constant volume flux was maintained throughout the entire time of experiments with adjustments in head of less than 0.4 cm of water. The soil columns were flushed free of  $\text{Cl}^-$  with tracer-free water before each experiment.

Miscible displacements of tracer and tracer-free waters were made by using the above-mentioned porous materials at different average flow velocities and different water contents.

**Results.** All curves shown in the figures to be presented in this paper are the result of measuring the chloride tracer concentration of effluent from the media. Consider Figure 2 for Yolo loam in which a 620-ml slug of water containing chloride tracer of concentration  $C_0$  is preceded



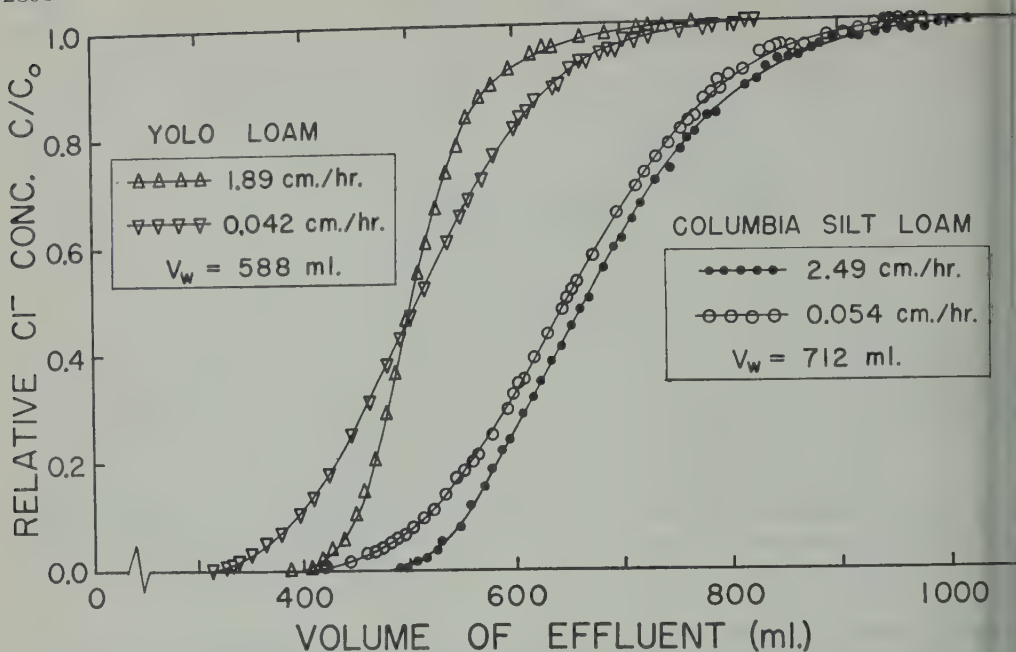
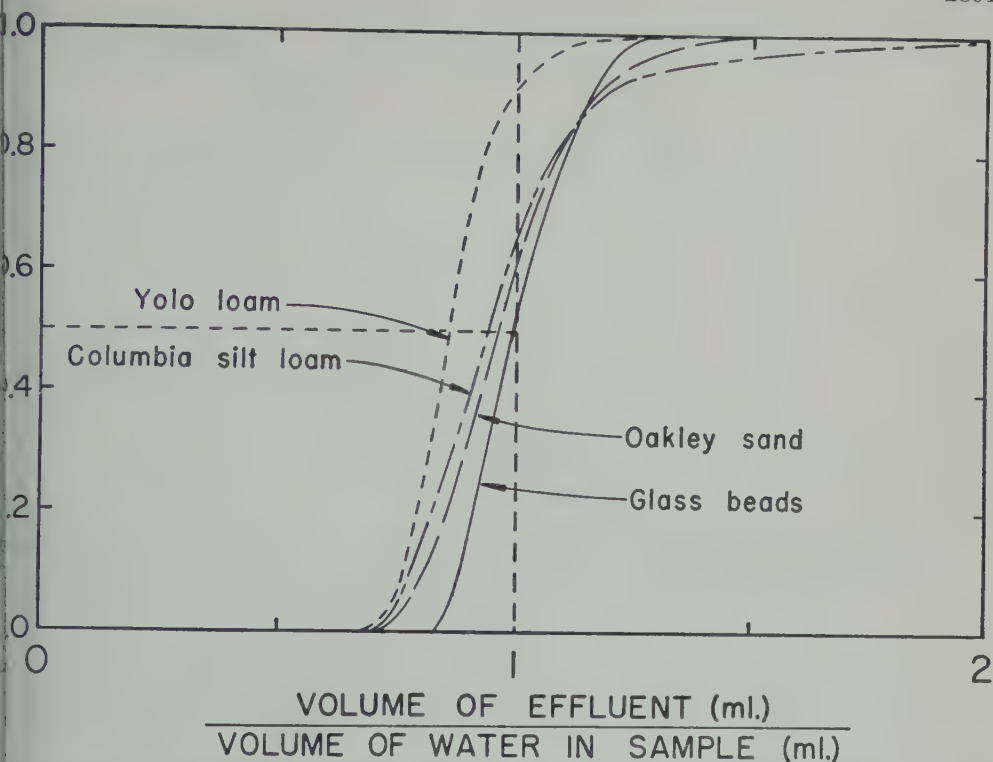


Fig. 3. Pairs of breakthrough curves using two average flow velocities for Yolo loam and Columbia silt loam. Both soils were water saturated.

and is followed by tracer-free water. If piston flow had occurred (i.e., no spreading of the tracer front) the  $\text{Cl}^-$  distribution or breakthrough curves would be represented by the broken vertical lines. If no ionic diffusion resulting from concentration gradients occurs in a medium having rather large microscopic flow velocities narrow in their distribution, a skewed sigmoidal breakthrough curve (Fig. 2) will pass through the point  $C/C_0 = 0.5$  at  $V_w$  (volumetric water capacity of column), and the areas above and below this point will be equal [Danckwerts, 1953]. If the flow velocity is everywhere zero and only ionic diffusion takes place at a tracer front within the medium, the resulting symmetrical sigmoid  $\text{Cl}^-$  distribution curves will pass through  $C/C_0 = 0.5$  at the original position of the tracer front. The curve in Figure 2 is skewed and sigmoid but does not pass through  $C/C_0 = 0.5$  at  $V_w$ . The general translation of the experimental breakthrough curve to the left of  $C/C_0 = 0.5$  at  $V_w$  is not expected when the foregoing remarks have been considered. This shift is caused by a significant fraction of the total pore volume not contributing to the volume

of effluent measured, a fact which will be fully dealt with later. The congruent nature of the experimental data for the first and second runs indicates the precision and accuracy of the method and proves that soil swelling or shrinkage was negligible during the experiment, a lapse of nearly 3 months.

The breakthrough curve presented in Figure 3 for Columbia silt loam at a flux of 0.054 cm/hr. is compared with that of a faster rate of 2.49 cm/hr. The same soil column was used for both flow velocities. The translation to the left at the smaller flow velocity as well as the slight clockwise rotation is caused by ionic diffusion combined with hydrodynamic dispersion during molecular displacement. In this respect, the breakthrough curve at the slower flow velocity might be expected to cross the breakthrough curve at the faster flow velocity [Rifai, Kaufman and Nielsen, 1956]. However, the flow which occurs in the smaller pores is sufficient to offset this effect. This is not true in the case of the Yolo loam also shown in Figure 3, for flow velocities of the same order of magnitude. The breakthrough curves do cross, suggesting an



Breakthrough curves for Yolo loam, Columbia silt loam, Oakley sand and glass beads under ed conditions. The average flow velocities were 1.89, 2.49, 0.30, and 1.77 cm/hr, respectively.

difference between the Yolo and Colum-  
s. It is apparent that for the Yolo soil a  
number of smaller pores act as static sinks  
hinder diffusion. As a result, the diffusion  
hinder from the larger pores having high  
concentrations into the smaller ones of  
concentration results in a less rapid break-  
a for the slower flow velocity and a  
transport volume before the maximum  
 $C/C_0 = 1$  is reached. In the Columbia,  
other hand, the smaller pores conduct  
more actively. This tends to mask the  
of any diffusion which has occurred.

are presented in Figure 4 to explain the  
ance of the early arrival of the measured  
through curves when compared with what  
normally have been expected had piston  
occurred. The abscissa in Figure 4 gives the  
of effluent divided by the volumetric  
capacity of a given porous material, i.e.  
number of pore volumes displaced. The dis-  
tance of the breakthrough curve to the left

of this vertical line or the area between the two  
lines is a relative measure of the volume of water  
not displaced but remaining within the sample.  
This area has been described previously as 'hold-  
back' [Danckwerts, 1953]. The magnitude of  
the translation would be expected to be associ-  
ated with soil texture and aggregation; sands  
with the smallest volume of water would be  
least displaced, and clays would have the great-  
est displacement. Figure 4 shows glass beads  
and Oakley sand to have the least holdback and  
Yolo loam to have the greatest. Because of the  
large flow velocities used for the four materials  
the effects of ionic diffusion on the translation  
of the breakthrough curves is not so obvious.

Equations based on the random-walk model  
and capillary models yield breakthrough curves  
which pass through the point  $C/C_0 = 0.5$  and  
pore volume = 1. Breakthrough curves reported  
in the literature have often been forced through  
this point for ease of comparison or for purposes  
of calculating the dispersion coefficient. Note in

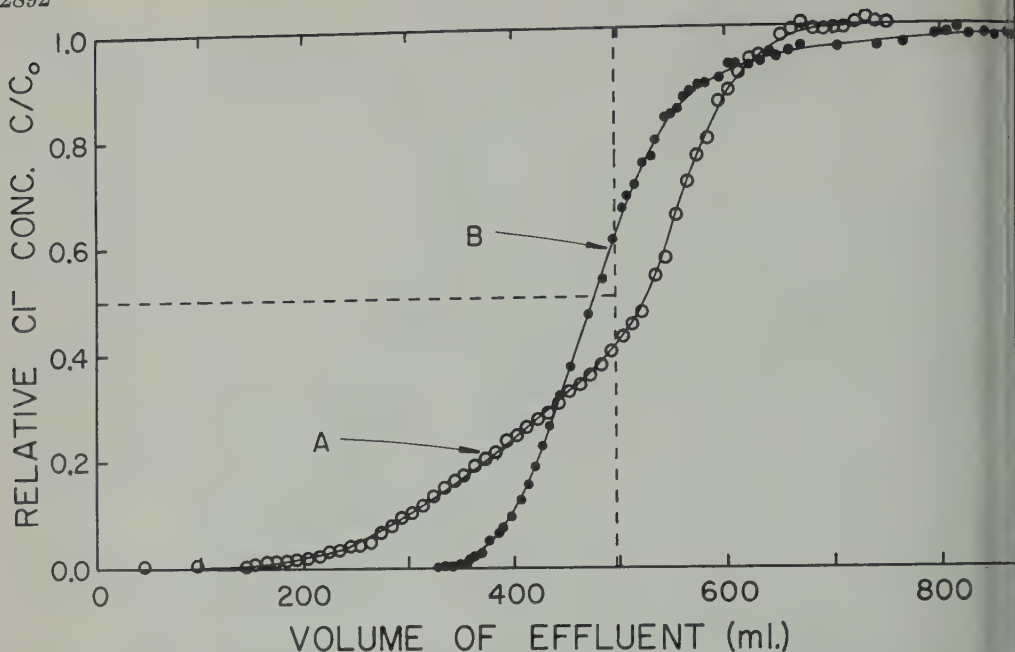


Fig. 5. Breakthrough curves obtained from the sample of Oakley sand under slight positive pressure (curve A) and under slight negative pressure (curve B). In both cases the flow velocities and water contents were identical.

Figure 4 that the breakthrough curve for the glass beads passes very close to  $C/C_0 = 0.5$  and 1 pore volume value and that the shape is characteristic of that predicted by the statistical models previously discussed. The medium dispersion coefficient  $D_m$  is  $0.20 \text{ cm}^2/\text{cm}$ . Apparently this porous material, composed of large uniform grains, satisfies the assumptions on which the above models are based. The other materials do not satisfy these assumptions since they are not completely without order and have a fraction of almost stagnant pores. The inadequacies of the theories will be more fully demonstrated later when a negative pressure is applied to the samples.

It has been shown in Figure 4 how a breakthrough curve may be translated to the left because of holdback. On the other hand, curves may be translated to the right for several reasons. The most obvious cause would be that of tracer exchange or precipitation within the soil column. Another might be electrical interactions between the tracer and media. We find that breakthrough curves will be translated to the right whenever flow takes place at the interface

of the soil and the soil container. This flow is not characteristic of that occurring in the bulk sample. Compare the two breakthrough curves presented in Figure 5 for the same soil sample for equal flow velocities and for identical initial chloride concentration in soil and solution. The use of small positive heads of water to control the flow rate yielded curve A while the use of small negative heads yielded curve B. The negative heads used did not unsaturate the column but merely drained about 5 ml of water associated with the lucite container-soil interface. For positive pressures, the contribution of breakthrough through these filled pores at the soil-soil-container boundary results in a curve not representative of the bulk soil. Hiby [see Safford 1960] has been concerned with this effect when experimentally determining the characteristic length of the pore canals of porous material. Consideration of this fact should be given in experiments involving miscible displacement studies in vertical columns under positive pressures.

Breakthrough curves obtained from unsaturated media yield additional information regarding

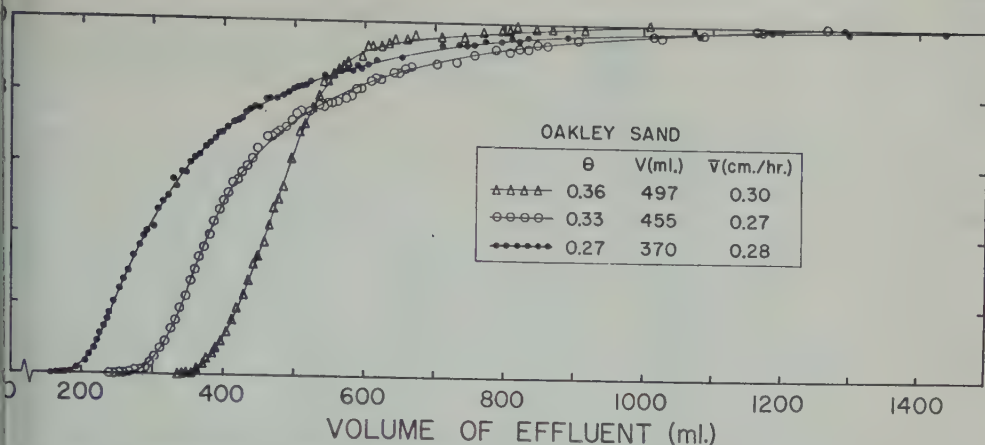


Fig. 6. Breakthrough curves for Oakley sand at three different water contents.

visible displacement not obtainable under the same conditions. Figure 6 shows breakthrough curves obtained for equal flow velocities in Oakley sand at three water contents. The greater the water content of the soil, the greater is the volume of effluent required to reach a maximum value (i.e.,  $C/C_0 = 1$ ). Desaturation eliminates larger flow channels and increases the volume of water that must be displaced. The volume of water in the sample which does not readily move, the stagnant water zones act as sinks for the diffusing substance. The relative increase in the volume of stagnant water caused by desaturation is more easily seen in Figure 7 where the

relative  $\text{Cl}^-$  concentration  $C/C_0$  is related to number of pore volumes leached through the column. Desaturating the soil has progressively shifted the breakthrough curves to the left at the initial breakthrough which may be explained by an incomplete displacement of one fluid by another. This characteristic shape resulting from incomplete mixing and 'dead water' was predicted by Danckwerts [1953]. Whereas diffusion is relatively unimportant in saturated sand, it becomes significant at the velocity used under desaturated conditions. This explains both the slow approach to a  $C/C_0$  value of 1.0 and the

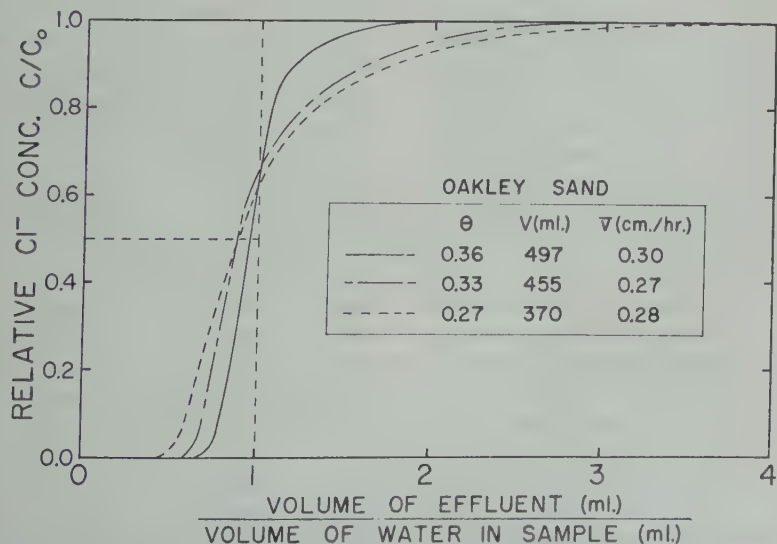


Fig. 7. Breakthrough curves for Oakley sand expressed as pore volumes displaced.



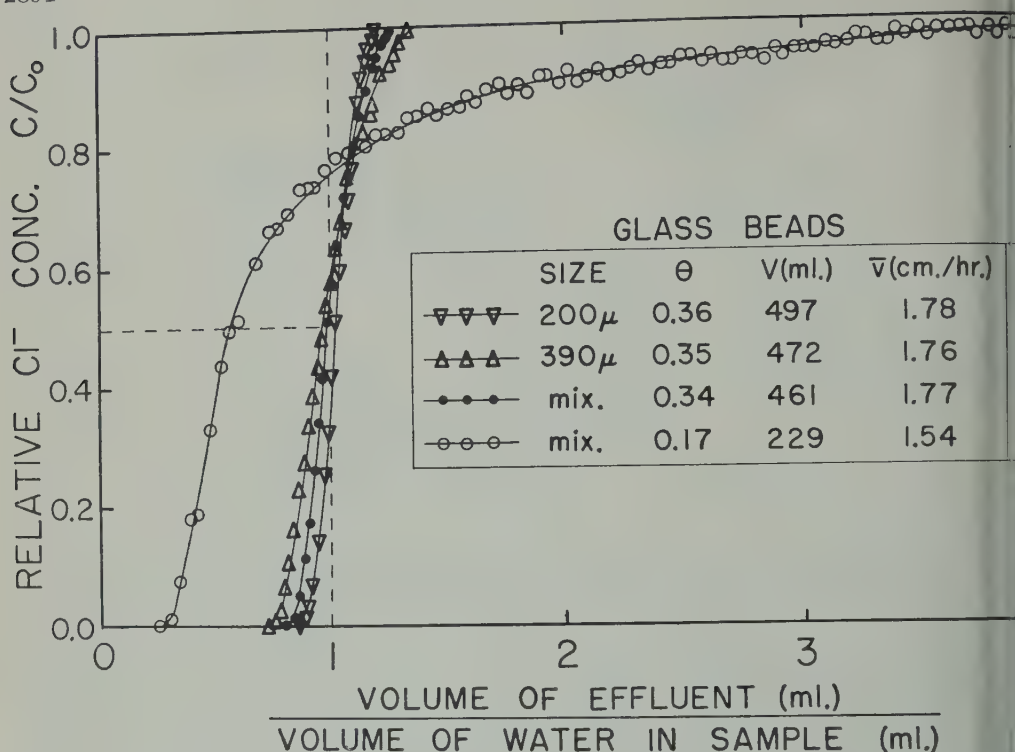


Fig. 8. Breakthrough curves for different sizes of glass beads under saturated and unsaturated conditions.

increase in skewness of the curves upon desaturation.

An experiment was run with negative pressure to determine whether the small amount of water associated with the emptied pores was active in transport—in other words, whether the voids created merely reduce the cross section for flow and act as solid grains of the material. Breakthrough curves were obtained by using columns of glass beads which are relatively inert compared with soils in order that the effects of the ionic charge of the chloride tracer be minimized. Glass beads having diameters of 200 and 390  $\mu$  were packed in a column after being mixed in proportions of 1 part to 6 on a weight basis. Another column was packed with only 200- $\mu$  beads and a third with only 390- $\mu$  beads. Breakthrough curves obtained under saturated conditions for flow velocities of the same order of magnitude are shown in Figure 8. These curves pass close to the point  $C/C_0 = 0.5$ , pore volume = 1, and have a normal sigmoidal shape, similar slopes,

and, in general, similar breakthrough characteristics.

A negative pressure applied to the mixture of glass beads decreased the volume of water in the sample from 461 to 232 ml. The breakthrough curve obtained at the reduced pressure is shown in Figure 8 by the early arrival of the tracer in the effluent translating the curve to the left of 1 pore volume. This translation results from flow occurring in the remaining water-filled pores whose gradient is in the direction of the principal axis of the column. The shape of the lower part of the breakthrough curve is similar to those obtained under saturated conditions. The asymptotic approach to  $C/C_0 = 1.0$  results from mixing due to extremely small microscopic flow velocities and ionic diffusion.

The water in equilibrium with the applied negative pressure resides primarily in the pores created by the 200- $\mu$  beads. Had the pressure produced by the applied pressure merely

the cross-sectional area for flow and acted as grain boundaries, a curve similar in shape to those obtained under saturated conditions would have been obtained.

The breakthrough curve for the unsaturated media is similar to those of the Oakley soil. As the magnitude of the holdback increases with decreasing water content and the slow approach to a maximum value of  $C/C_0 = 1$  with an unsaturated sample is similar to that of a saturated sample. Both of the above facts provide further evidence of incomplete displacement of one fluid by another occurring in saturated samples prior to breakthrough in sequences of the smaller water-filled pores. As a consequence of the above findings in unsaturated materials, it may be considered that the slow approach to a maximum  $C/C_0 = 1$  under saturated conditions is caused by the extremely large contribution of  $Cl^-$  concentration from the larger pore sequences having the slowest flow velocities. Even if the concentration,  $C/C_0$ , is not measurably different under saturated conditions, it should not be concluded that the concentration of tracer in the column is  $C_0$ . Eventually, however, all the tracer in the column will approach  $C_0$  by tracer breakthrough.

Further work is now in progress to investigate adsorption and diffusion in relation to exchange adsorption, in saturated and unsaturated porous materials.

#### REFERENCES

Arfken, W. J., Dispersion of soluble matter in slowly moving fluids, Ph.D. Thesis, Harvard University, Cambridge, Mass., 92 pp., May 1955.  
 Cernik, P. V., Continuous flow systems; distribution of residence times, *Chem. Eng. Sci.*, **2**, 1953.  
 Day, P. R., Dispersion of a moving salt-water front advancing through saturated sand,

*Trans. Am. Geophys. Union*, **37**, 595-601, 1956.  
 Day, P. R., and W. M. Forsythe, Hydrodynamic dispersion of solutes in the soil moisture stream, *Proc. Soil Sci. Soc. Am.*, **21**, 477-480, 1957.  
 de Josselin de Jong, G., Longitudinal transverse diffusion in granular deposits, *Trans. Am. Geophys. Union*, **39**, 67-74, 1958.  
 Handy, L. L., An evaluation of diffusion effects in miscible displacement, *J. Petrol. Technol.*, 61-63, 1959.  
 Kaufman, W. J., and G. T. Orlob, An evaluation of ground-water tracers, *Trans. Am. Geophys. Union*, **37**, 297-306, 1956.  
 Orlob, G. T., and J. N. Radhakrishna, The effects of entrapped gases on the hydraulic characteristics of porous media, *Trans. Am. Geophys. Union*, **39**, 648-659, 1958.  
 Rifai, M. N. E., W. J. Kaufman, and D. K. Todd, Dispersion phenomenon in laminar flow through porous media, *Progr. Rept. 2*, Canal Seepage Research, Univ. Calif., Berkeley, 157 pp. 1956.  
 Saffman, P. G., Dispersion due to molecular diffusion and macroscopic mixing in flow through a network of capillaries, *J. Fluid Mech.* **7**(2), 194-208, 1960.  
 Scheidegger, A. E., Statistical hydrodynamics in porous media, *J. Appl. Phys.*, **25**, 994-1001, 1954.  
 Scheidegger, A. E., On the theory of flow of miscible phases in porous media, *Extrait des Comptes Rendus et Rapports, Assemblee Generale de Toronto*, Tome II, pp. 236-242 Gentbrugge, 1958.  
 Slichter, C. S., Field measurements of the rate of movement of underground waters, *U. S. Geol. Survey, Water Supply and Irrigation Paper* **140**, 122 pp., 1905.  
 Taylor, G. I., Conditions under which dispersion of a solute in a stream of solvent can be used to measure molecular diffusion, *Proc. Roy. Soc. London, A*, **225**, 473-477, 1954.  
 von Rosenberg, D. V., On the mechanics of steady-state single phase fluid displacement from porous media, *A. I. Ch. E. Journal*, **2**, 55-58, 1956.  
 Yuhara, K., A model experiment of saline diffusion in flow of water through porous media, *Geophysics*, **9**, 127-134, 1954.

(Manuscript received June 6, 1960.)



## Radioactivity Associated with Underground Nuclear Explosions

ROGER E. BATZEL

*University of California  
Lawrence Radiation Laboratory  
Livermore, California*

**Abstract.** The detonation of a contained or partially contained nuclear explosion is accompanied by the deposition of a large fraction of the energy in the form of high-temperature, high-pressure regions. The nature of the surrounding medium, the time-temperature history, and the time of cavity collapse or venting determine the extent to which undesirable nuclides such as  $\text{Sr}^{90}$  and  $\text{Cs}^{137}$  will appear outside a fused, relatively insoluble matrix and be available to ground water or to the atmosphere. The movement of these undesirable radioactive nuclides relative to the ground-water movement can be predicted on the basis of measured  $K_D$ 's (distribution coefficients) for the radioactivities in the medium.

The induced radioactivities are a 20 to 25 per cent contribution to the fission product radioactivity for times of the order of 1 day, and a 1 per cent contribution for times equal to about a week. The contribution decreases to 0.1 per cent after about 45 days, but later, because of the  $\text{Co}^{60}$  present, it increases to about 2 per cent for a period of 3 to 15 years.

This paper was issued originally by the Atomic Energy Commission as an internal report with the title, as UCRL-5623, June 23, 1959. At that time a fair amount of the material published here in the hope that it will be of use to those involved in the evaluation of potential hazards associated with underground nuclear explosions.

When a nuclear explosion is fired underground, subsequent to the immediate release of energy resulting from the fission or thermonuclear reactions the energy of the explosion is deposited in the immediately surrounding cavity at temperatures of the order of a few million degrees [Johnson and Violet, 1958; Johnson, 1959; and Violet, 1959]. At these temperatures the chemical behavior of the fission products is unimportant, as the matter is converted to a gas of ionized atoms and electrons.

As the shock progresses into the wall of the cavity the first few feet of the surrounding material is vaporized (about 1 meter for a 20-kiloton equivalent explosion) and is mixed with the gases containing the fission products. As the gas cools to a temperature of about 1000°C, condensed phases can begin to form, and a condensed phase is then deposited as a layer on the molten layer on the wall of the

cavity. The major part of the condensation is probably completed in times of the order of several hundred milliseconds. Water vapor and noncondensable gases, including rare gases from the fission process and the more volatile fission products and volatile constituents of the medium, will, however, be left in the gaseous state.

The details of the above phenomena depend on the time-temperature history, the degree of containment and associated collapse phenomena, and the nature of the medium in which the explosion takes place. If the cavity collapses or vents at a time when the pressure and temperatures are still high, the major fraction of the material in the gas phase will escape from the immediate highly radioactive molten zone.

**Fission process.** The prompt energy produced in fission, including kinetic energy of the fission fragments, prompt  $\gamma$  rays, and neutrons, is 179 Mev [Gunn, Hicks, Levy, and Stevenson, 1957]. A kiloton equivalent is defined as  $1.0 \times 10^{12}$  calories or  $4.185 \times 10^{10}$  ergs, and, on the basis of 179 Mev per fission, there are  $1.46 \times 10^{28}$  fissions per kiloton. In addition to the prompt energy from fission, an average of about 22 Mev of additional energy results from decay of the fission products. About 50 per cent appears in the form of  $\gamma$  rays, about 17 per cent as  $\beta$ -decay energy, and the other 33 per cent as neutrinos



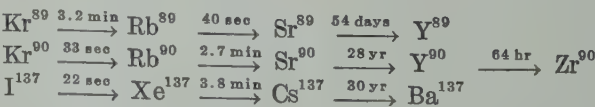
associated with the  $\beta$ -decay process. About 7 Mev of the  $(\beta + \gamma)$ -decay energy appears in the first 20 minutes after detonation.

The fission of about 55 grams of  $\text{Pu}^{239}$ ,  $\text{U}^{235}$ , or  $\text{U}^{238}$  yields an energy equivalent of 1 kiloton. For example a 1-megaton nuclear explosion with a fission-to-fusion ratio of 0.05 has a fission yield of 50 kilotons, with its associated 2750 grams of fission products. These 2750 grams of fission products contain about 50 grams of  $\text{Sr}^{90}$  and are the equivalent of the waste fission products from the operation of a 1000-megawatt reactor for approximately  $2\frac{3}{4}$  days.

The detonation of a nuclear explosion produces fission products in proportion of the fission yield of the explosion, and these fission products are distributed according to the typical fission-yield curve [Coryell and Sugarman, 1951]. The general characteristics of this fission-yield curve are the same for the fission of  $\text{U}^{238}$ ,  $\text{U}^{235}$ , and  $\text{Pu}^{239}$ .

It is necessary to consider the elemental form of the fission products at the time of cavity venting in order to estimate the fraction of the radioactivity which is in the gaseous state. Table 1 illustrates the contribution of the elements as percentages of the total radioactivity at various times after the detonation. These data are taken from the calculations of Bolles and Ballou [1956] and apply to products of the thermal fission of  $\text{U}^{235}$ .

It is obvious from Table 1 that a large fraction of the radioactivity exists in the form of noble gases or somewhat volatile elements on time scales which should apply to situations in which cavity venting would occur. Nuclides in the form of the noble gases will not condense until decay to another element has occurred, and the more volatile elements will not condense until the temperature is lowered or a decay to a more refractory element occurs. The following are typical examples of fission product decay chains:



As an example for the mass-90 chain which is formed in fission, most of the material which finally exists as  $\text{Sr}^{90}$  is produced in the primary fission as  $\text{Kr}^{90}$  (about 80 per cent with smaller amounts formed directly as  $\text{Br}^{90}$  or  $\text{Rb}^{90}$ ). Thus,

TABLE 1. Percentage Contribution of Activity of Elements to Total Radioactivity\*

| Element | 10 sec. | 30 sec. | 1 min. | 5 min. | 30 min. |
|---------|---------|---------|--------|--------|---------|
| Kr      | 3.8     | 5       | 5.8    | 4.3    | 2.5     |
| Xe      | 2.6     | 3.8     | 4.6    | 9      | 5       |
| Y       | 11      | 9       | 8      | 4.6    | 7       |
| Zr      | 12      | 7       | 3.5    | ...    | ...     |
| Ba      | 7       | 2.6     | 2.4    | 8      | 9       |
| Sr      | 6       | 4.0     | 5      | 11     | 4.0     |
| Cs      | 7       | 7       | 7      | 8      | 8       |
| Rb      | 7       | 8       | 8      | 11     | 7       |
| La      | 9       | 8       | 4.0    | 2.8    | 9       |
| Nb      | 10      | 12      | 11     | 8      | 5       |
| Ce      | 4.7     | 3.0     | 1.0    | 2.5    | 5       |
| Br      | 3.3     | 5       | 6      | 2.2    | ...     |
| I       | 4.0     | 6       | 7      | 3.5    | 3.5     |
| Mo      | 3.0     | 6       | 6      | 5      | 7       |
| Te      | 2.7     | 2.5     | 2.0    | 3.5    | 9       |
| Se      | 1.5     | 1.0     | ...    | ...    | ...     |
| Pr      | 1.7     | 3.7     | 5      | 4.6    | 3.7     |
| Sn      | 1.1     | ...     | ...    | ...    | ...     |
| Tc      | 1.2     | 4.0     | 8      | 5      | 9       |
| Sb      | 1.0     | 2.3     | 3.7    | 5      | 2.3     |
| Nd      | ...     | ...     | ...    | ...    | ...     |

\* Radioactive nuclides contributing less than 1 per cent of the total radioactivity are not included.

if the cavity vents in times comparable to the half-life of the  $\text{Kr}^{90}$ , a major fraction of the  $\text{Sr}^{90}$  will escape along with the other gases.

Typical measurements made on debris from several of the explosions at the Nevada Test Site are [Goeckermann, 1960]:

|                  |   |
|------------------|---|
| $\text{Sr}^{89}$ | 3 to 10 per cent in the major part of the fused tuff, enriched in the rubble by severalfold and in the vented debris by nearly an order of magnitude. |
| $\text{Sr}^{90}$ | 20 to 30 per cent in the fused tuff, vented debris enriched by fivefold, local fallout less enriched.   |
| $\text{Y}^{91}$  | 30 to 50 per cent in the fused tuff.  |

|                  |  |
|------------------|--|
| $\text{Zr}^{95}$ | vented debris and rubble enriched severalfold.<br>70 to 100 per cent in the fused tuff, rubble enriched. |
|------------------|--|

- 20 to 40 per cent in the fused tuff, vented material enriched by several fold.  
 30 to 60 per cent in fused tuff decreasing as yield increases, vented debris enriched by several fold.  
 ~100 per cent in the fused tuff.  
 ~100 per cent in the fused tuff, precursors are rare earths.

*completely contained explosions.* On the basis of the discussion, even with the formation of an insoluble glassy matrix in which the major portion of the fission products is incorporated, the fraction of the  $\text{Sr}^{90}$  and  $\text{Cs}^{137}$ —that fraction which exists as a rare gas at the time of the collapse—will not be incorporated but will be distributed with the other gases at the time of cavity collapse. Thus appreciable amounts of  $\text{Sr}^{90}$  and other nuclides with rare earth precursors or nuclides of volatile elements will be expected to be deposited at appreciable distances from the highly radioactive zone and will be distributed on the surfaces of the broken material as decay products of the rare gas precursors.

*Partially contained explosions.* Depending on the time and extent of venting of a partially contained explosion, the fraction of radioactive material which does vent to the atmosphere should be small. In the rare gas and more volatile fission products,  $\text{Sr}^{90}$  and  $\text{Cs}^{137}$  should appear enriched with respect to the nuclides which do not have rare earth precursors existing as rare gases at the time the radioactive debris is vented to the atmosphere. For example, samples of the vented debris from the Blanca (20-kiloton equivalent) detonation, which vented approximately 0.3 to 0.5 per cent of the total fission activity, was enriched about fivefold in  $\text{Sr}^{90}$ , indicating that about 0.5 per cent of the total  $\text{Sr}^{90}$  formed in the explosion was actually released to the atmosphere. As the rare gases decay to such elemental products as  $\text{Rb}^{90}$  and subsequently to  $\text{Sr}^{90}$ , the  $\text{Sr}^{90}$  would be expected to deposit rapidly upon the surface of the large amount of particulate material which is blown out with the vented gas and be scavenged from the atmosphere as the particulate matter falls out.

*Ground-water contamination.* From the discussion it follows that the formation of an insoluble glassy material during the explosion

does not guarantee the trapping of all the undesirable radioactive nuclides. In fact, large fractions of two of the more undesirable nuclides from a biological point of view,  $\text{Sr}^{90}$  and  $\text{Cs}^{137}$ , will appear outside the fused material and in the broken material, and, regardless of the existence of fused, glassy, insoluble material in which the majority of the other fission products are contained, these nuclides will be accessible to ground water.

Higgins [1959], in his paper on ground water, discussed the fact that ions such as  $\text{Sr}^{90}$  and  $\text{Cs}^{137}$  are selectively adsorbed from ground water by most naturally occurring minerals and that these ions distribute themselves between water and minerals according to the following law:

$$K_D = A_S M_w / A_w M_S$$

where  $K_D$  is the distribution coefficient,  $A_S$  and  $A_w$  are the activities in the mineral and water, respectively, and  $M_w$  and  $M_S$  are the masses of water and mineral, respectively, in equilibrium. Given  $K_D$  of a nuclide, it is possible to calculate the average flow rate of the activity as a function of the ground-water flow rate.

$$F_A = F_w \frac{1}{1 + K_D \rho}$$

where  $\rho$  is the ratio of the mass of mineral per unit volume to the mass of water per unit volume.  $\rho = 10$  represents the typical situation in which most ground water occurs. The  $K_D$ 's for naturally occurring minerals range from minima of about 40 (limestone) for  $\text{Sr}^{90}$  to 100,000 for  $\text{Ce}^{144}$ . With a typical flow rate for ground water of 1 meter per day, for example, and assuming  $K_D = 300$  and  $\rho = 1$ , the average flow rate of  $\text{Sr}^{90}$  in the medium would be

$$F_A = 1 \text{ meter/day} \times \frac{1}{1 + 300} = 0.003 \text{ meter/day}$$

so that in a time equal to the half-life of  $\text{Sr}^{90}$  (28 years) the  $\text{Sr}^{90}$  activity would have moved about 30 meters. See the Appendix for an example of the effect of the distribution coefficient on the concentration of  $\text{Sr}^{90}$  in ground water.

On the basis of the above considerations it is possible to predict the fate of these radioactive nuclides in the ground water by having a de-

TABLE 2. Composition of a Typical Medium

| Element | Abundance,<br>wt. per cent |
|---------|----------------------------|
| Si      | 50                         |
| Al      | 14.5                       |
| Fe      | 9                          |
| Ca      | 6.4                        |
| H       | 5.1                        |
| Na      | 4.9                        |
| K       | 4.7                        |
| Mg      | 3.7                        |
| Ti      | 0.6                        |
| P       | 0.18                       |
| Mn      | 0.18                       |
| Co      | 0.0042                     |

tailed knowledge of the hydrology of an area and by making laboratory measurements on minerals which will be in contact with the ground water of the  $K_D$ 's of the biologically undesirable nuclides.

*Induced radioactivity.* On the basis of the assumption that  $\sim 1 \times 10^{23}$  neutrons per kiloton equivalent of yield are emitted from a nuclear explosion and that these  $1 \times 10^{23}$  neutrons are captured in the surrounding medium, it is possible to estimate the levels of radioactivity induced in a typical medium surrounding a nuclear explosion. The major fraction of the induced radioactivity would be expected to occur in the first meter of medium and to be included in the initially vaporized material.

Table 2 gives the typical medium composition chosen [Mason, 1952]. Oxygen is not included

TABLE 3. Radioactive Nuclides Produced

| Element                         | Nuclide<br>Produced | Half-<br>Life* | $\gamma^*$<br>Energy,<br>Mev | Percent-<br>age of<br>Neutron<br>Captures<br>Lead-<br>ing to<br>Nuclide |
|---------------------------------|---------------------|----------------|------------------------------|---|
| Na                              | Na <sup>24</sup>    | 15 hours       | 4.0                          | 3.7   |
| Al                              | Al <sup>28</sup>    | 2.3 min.       | 1.8                          | 2.6   |
| Mn                              | Mn <sup>56</sup>    | 2.6 hours      | 1.8                          | 1.3   |
| Fe<br>(0.33% Fe <sup>58</sup> ) | Fe <sup>59</sup>    | 45 days        | 1.3                          | 0.015   |
| Co                              | Co <sup>60</sup>    | 5.2 years      | 2.5                          | 0.09  |

\* Strominger, Hollander, and Seaborg, [1958].

TABLE 4. Curies per Kiloton Equivalent Nuclides Produced

| Nuclide          | Curies            | (Curies) $\times$ ( $\gamma$ Energy) |
|------------------|-------------------|--------------------------------------|
| Na <sup>24</sup> | $1.3 \times 10^6$ | $5.2 \times 10^6$                    |
| Al <sup>28</sup> | $3.5 \times 10^8$ | $6.3 \times 10^8$                    |
| Mn <sup>56</sup> | $2.6 \times 10^8$ | $4.7 \times 10^8$                    |
| Fe <sup>59</sup> | $\sim 10^3$       | $\sim 1.3 \times 10^3$               |
| Co <sup>60</sup> | $\sim 10$         | $\sim 25$                            |

because it absorbs substantially less than 1 per cent of the neutrons. The composition of the water content of the typical medium is 20 per cent by weight.

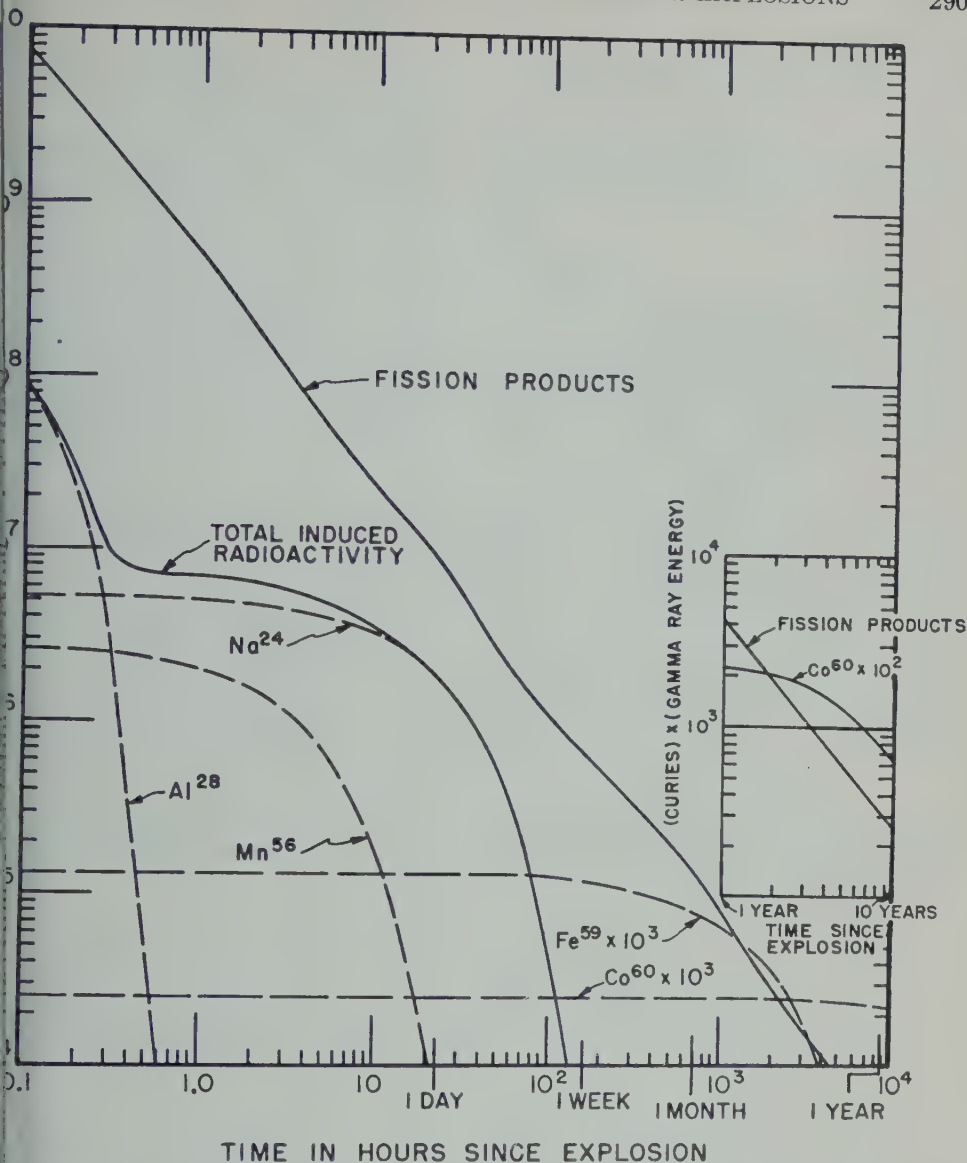
The neutrons captured in the typical medium were distributed according to the atom per cent of the element present and the thermal neutron capture cross sections for the element [Hussey and Schwartz, 1958]. If the water content is 20 per cent by weight, most of the neutrons are thermalized before capture occurs, and about 60 per cent of the neutrons are captured by the hydrogen of the water in this typical medium. The important radioactive nuclides produced which include  $\gamma$  rays in their decay schemes, and the percentage of the neutron capture leading to these nuclides are shown in Table 3.

The number of curies per kiloton equivalent of these nuclides estimated to be produced are listed in Table 4.

The number of (curies)  $\times$  ( $\gamma$ -ray energy) is a measure of the relative biological importance of the  $\gamma$  rays as a function of time. Figure 1 is a comparison of the radioactivity of the induced nuclides with the radioactivity of the fission products from 1 kiloton equivalent of fission yield. The induced radioactivities are a 20 per cent contribution to the fission product radioactivity at times of the order of 1 day, a 1 per cent contribution for times of about 1 week. This contribution decreases to 0.1 per cent in about 45 days, but it increases later, due to the presence of Co<sup>60</sup>, to about 2 per cent after a period of 3 to 15 years.

Not included in the above discussion was the production of nuclides which do not produce energetic  $\gamma$  rays in their decay and would not contribute to an external  $\gamma$ -radiation field which may constitute potential biological problems. For example, Ca<sup>45</sup> (half-life 164





Comparison of induced radioactivity and fission product radioactivity from a one kiloton equivalent nuclear explosion.

rays by negative  $\beta$  emission, is proportional to the extent of about 200 curies per equivalent by the emergent neutrons and is well mixed and diluted with a minimum of 500 tons of the medium.

#### APPENDIX

*Rate of Distribution of Sr<sup>90</sup> from a Nuclear Explosion.* If it is assumed that all the Sr<sup>90</sup>

escapes from the cavity and is distributed in an amount of broken medium equivalent to the fused material, as observed in Nevada tuff (~ 500 tons of fused material per kiloton), the concentration of Sr<sup>90</sup> in the broken medium is (1 gram or ~150 curies of Sr<sup>90</sup> per kiloton equivalent of fission)

$$\frac{150 \text{ curies}}{500 \text{ tons} \times 10^6 \text{ g/ton}} \approx 3 \times 10^{-7} \text{ curies/g}$$



The actual situation is probably better represented by the distribution of the 150 curies of  $\text{Sr}^{90}$  in some 5000 to 50,000 tons of broken medium.

If a typical  $K_d$  of 300 is used, the concentration of  $\text{Sr}^{90}$  in ground water in contact with the medium is

$$\frac{3 \times 10^{-7} \text{ curies/g}}{300} = 1 \times 10^{-9} \text{ curies/g}$$

of water. After the 1 gram of ground water has equilibrated with an uncontaminated gram of the medium, the concentration of  $\text{Sr}^{90}$  in the ground water is

$$\frac{1.0 \times 10^{-9} \text{ curies/g}}{300}$$

or  $\sim 3.3$  micromicrocuries/g of water, or about 10 times the tolerance level. It is clear that equilibration of the 1 gram of the water with another gram of medium will leave the concentration of  $\text{Sr}^{90}$  below the tolerance level.

*Acknowledgment.* This work was performed under the auspices of the U. S. Atomic Energy Commission.

#### REFERENCES

Bolles, R. C., and N. E. Ballou, Calculated activities and abundances of  $\text{U}^{235}$  fission products,

- USNRDL-456*, U. S. Naval Radiological Defense Laboratory, San Francisco, 1956.  
 Coryell, C. D., and N. Sugarman, *Radiochemical Studies: The Fission Products, Book 3*, McGraw-Hill Book Co., New York, 1951.  
 Goeckermann, R. H., The disposition of radioactivity in underground nuclear explosions, *UCRL-5675*, Part I, Lawrence Radiation Laboratory, University of California, 1960.  
 Gunn, S. R., H. G. Hicks, H. B. Levy, and J. Stevenson, A calorimetric determination of average total kinetic energy of fragments in fission of  $\text{U}^{235}$ , *Phys. Rev.*, **107**, 1642-1645, 1959.  
 Higgins, G. H., Evaluation of the ground water contamination hazard from underground nuclear explosions, *J. Geophys. Research*, **64**, 1509-1519, 1959.  
 Hughes, D. J., and R. B. Schwartz, Neutron cross sections, *BNL-325*, Brookhaven National Laboratory, 1958.  
 Johnson, G. W., G. H. Higgins, and C. E. Violet, Underground nuclear detonations, *J. Geophys. Research*, **64**, 1457-1470, 1959.  
 Johnson, G. W., and C. E. Violet, Phenomena of contained nuclear explosions, *UCRL-1000*, Rev. I; Lawrence Radiation Laboratory, University of California, 1958.  
 Mason, Brian, *Principles of Geochemistry*, Wiley & Sons, New York, 1952.  
 Strominger, D., J. M. Hollander, and G. T. Seaborg, Table of isotopes, *Revs. Modern Physics*, **28**, 585-904, 1958.

(Manuscript received May 23, 1960; revised June 20, 1960.)

# Natural Radiocarbon in the Atlantic Ocean<sup>1</sup>

WALLACE S. BROECKER, ROBERT GERARD, MAURICE EWING, AND BRUCE C. HEEZEN

*Lamont Geological Observatory  
Columbia University  
Palisades, New York*

**Abstract.** By the use of suitable chemical and radiometric techniques the natural radiocarbon concentration in the dissolved bicarbonate of 135 samples representing the major water masses of the Atlantic Ocean has been determined with a precision ranging from 0.5 to 1.3 per cent. Whereas the results from a given water mass exhibit a standard deviation only slightly exceeding that predicted from the experimental error alone, measurable differences exist between the major water masses, the total range in  $C^{14}/C^{12}$  ratio being about 10 per cent. Corrections for the bomb-produced  $C^{14}$  effect and the industrial  $CO_2$  effect have been applied where necessary.

The surface water  $C^{14}/C^{12}$  ratios show a progressive increase from south to north, ranging from 120 per mil lower than the preindustrial atmospheric value in the Antarctic to 50 per mil lower in the North Atlantic. Deep water masses originating in the high latitudes of the northern hemisphere have consistently lower  $C^{14}/C^{12}$  ratios than those originating in the high latitudes of the northern hemisphere. A layer of water of high  $C^{14}/C^{12}$  ratio found at depths between 1200 and 2400 meters in the western North Atlantic may well represent a wedge of young water penetrating the older North Atlantic deep water. Bottom water in the eastern basin of the North Atlantic has a 20 per mil lower  $C^{14}/C^{12}$  ratio than the corresponding water in the western basin.

According to a steady-state circulation model, most of the water below 600 meters in the North Atlantic remains at depth for an average of 650 years. Corresponding residence times for water masses of Antarctic origin are less than 350 years.

A circulation model explaining the prominent features of the  $C^{14}$  distributions in the atmosphere-ocean system is based on a south to north transport of water along the surface of the Atlantic Ocean, with a return flow at depth. The Atlantic and Pacific communicate through the Antarctic. On the basis of this model, despite the lower  $\Delta C^{14}$  values, the mean residence times of water in the deep reservoirs of the Pacific may not exceed those for the deep Atlantic more than 30 per cent.

Although results of  $C^{14}$  analyses on tree rings suggest that the oceans are reasonably close to steady state, the possibility of nonsteady-state circulation must be considered. It is shown that the present  $C^{14}$  distribution in the oceans could be achieved through the storage of  $C^{14}$  in the atmosphere and surface oceans during a relatively short period of greatly restricted bottom water formation. If nonequilibrium effects are important the residence times computed from the steady-state model could be considerably in error.

**Introduction.** The absolute rate of overturn in the deep ocean is of considerable interest not only to the oceanographer but also to those engaged in other branches of earth science. For example, because of the large heat capacity of the deep sea, its rate of interaction with the surface ocean and atmosphere is of great interest to the meteorologist and the biologist. The biologist interested in the distribution of nutrients in the ocean also needs esti-

mates of residence times in the deep sea. More recently those concerned with the problem of radioactivity produced by bombs and reactors have looked to the oceans as a diluent for these products. Again, mixing rates are of prime importance.

Although considerable progress has been made in understanding the patterns and rates of mixing in the oceans by traditional studies of density distribution, the problem is an exceedingly complex one, requiring as many independent approaches as possible. One such approach, the study of the distribution of  $C^{14}$  in the oceans,

was first applied in 1950 by Ewing and Kulp. The original  $C^{14}$  measurements reported by Kulp, Tryon, Eckelmann, and Snell [1952] were later found to be somewhat in error owing to contamination in the ascarite absorbent used in shipboard processing and to the inherent problems of the black carbon assay system. These results, however, showed that the residence time for deep waters in the Atlantic Ocean was less than 2000 years. In 1955 the present study was initiated at Lamont. A new water processing system and the more precise gas counting assay technique [Broecker, Tucek, and Olson, 1959] were used. Preliminary results of this investigation have been given by Broecker [1957] and Broecker, Ewing, Gerard, Heezen, and Kulp [1958]. Other laboratories are also currently carrying out similar studies. Results of these investigations have been published by Rafter [1955], Rafter and Fergusson [1958], Brodie and Burling [1958], Garner [1958], Rubin and Alexander [1958], Fonselius and Ostlund [1959], and Suess, Rakestraw, and Oeschger [1959]. The work reported here was carried out as a joint project between the geophysics and geochemistry groups at Lamont, the former being responsible for the design of the water sampler and the collection and shipboard chemical processing of the water samples and the latter for the design of the shipboard processing system and the laboratory analysis of samples.

The object of the present research was to determine the geographic and depth variation in the Atlantic Ocean of the ratio of  $C^{14}$  (a radioactive isotope of 5570 years half-life), to  $C^{12}$  (the most abundant stable isotope of carbon). If it is assumed that factors other than the radioactive decay producing these variations could be recognized and accounted for, then the residual differences should be related to mixing rates.

*Experimental techniques.* The procedures for sampling sea water for  $C^{14}$  have been designed to collect large-volume samples from accurately known depths in the ocean, to extract the inorganic carbon in the form of  $CO_2$ , and to return that portion, free from contamination, to the laboratory for  $C^{14}/C^{12}$  measurements.

The total concentration of inorganic  $CO_2$  in all forms in ocean water is about 0.025 grams of carbon per liter [Sverdrup, Johnson, and Fleming, 1942]. The volume of the Lamont counters

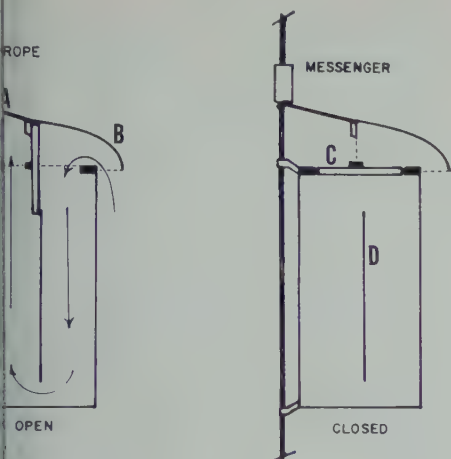
for measuring  $C^{14}$  content is 5 liters. Five grams of carbon in form of  $CO_2$  gas is required to produce 2 atmospheres of pressure in the counters. On the basis of this requirement, a minimum of 200 liters of sea water is needed for a sample. During the period 1955-1959 400-l samples were collected to allow for variations in the efficiency of the recovery system. Currently a 220 liter sampler is sufficient because of improvements in recovery techniques.

Both types of sampler were designed to be lowered into the ocean by a winch using  $\frac{1}{2}$ -inch wire rope to which the sampler is securely attached. The lowering is accomplished with a 1500-pound piston-coring device fixed to the wire rope below the water sampler as a weight. The earlier 400-liter sampler consisted of a steel barrel with two round, 10-inch-diameter ports at the top fitted with hinged doors which opened and closed in unison. As the sampler was lowered through the water a hood-like structure directed the flow of water into one of the ports while a partition inside the sampler further directed the flow to the bottom of the barrel, from whence the water moved upward out of the other open port. Dye experiments have shown that, with this method of flushing, less than 0.1 per cent of an original water volume remains in the sampler after it is lowered 100 meters. The same principle is used in the present sampler (Fig. 1) but it has only a single door which when open divides the single port into two and when closed seals the port with a ring seal.

When the sampler is lowered to the desired depth, a 'messenger' in the form of a 5-pound weight is dropped down the wire from the surface. The messenger strikes a release on the sampler which closes the doors. In order to interpret the results with respect to water mass in the ocean, precise knowledge of depth at sampling is necessary. To afford additional control to that provided by length and angle of wire and by sounding, protected and unprotected reversing thermometers are used inside the sampler. When the door closes, the thermometer reverses, thereby indicating the depth and temperature of the sample. A small Bourdon recorder is also used to indicate the depth at closing of the sampler door.

When the sampler is brought up to the





## LARGE-VOLUME WATER SAMPLER

1. Idealized diagram of 220-liter water sampler presently in use. The drawing on the left shows the open position. Arm *A* is the trigger arm which releases the closing mechanism. Hood *B* directs the flow of water into the barrel while lowering the drawing on the right shows the sampler lowered position. Round plastic door *C* has an air seal and closes when messenger strikes trigger arm *A*. Partition *D* channels water to both sides of sampler to aid circulation while lowering.

Water samples are taken for determination of pH and dissolved oxygen. A submersible pump is then placed in the bottom of the barrel and the water is pumped through a hose to a surface deck tank where processing for removal of CO<sub>2</sub> is carried out.

The water in the processing tank is acidified with H<sub>2</sub>SO<sub>4</sub> and bubbled in a closed system with either air or tank N<sub>2</sub>. In the present system, CO<sub>2</sub> recovery is achieved by using a pump to circulate the water within the tank and strip it into the gas phase. The CO<sub>2</sub> released from the gas phase is absorbed in 5*N* KOH held in a series of water bubblers operated in series. Recovery is about 90 per cent of the CO<sub>2</sub> from a 220-liter sample. The recovery is checked by means of a titration technique. The KOH is returned to the laboratory, where it is acidified again and the CO<sub>2</sub> is released. The CO<sub>2</sub> is purified by absorption and desorption in NaOH contained in quartz tubes and is then assayed for radioactivity in a 5-liter proportional counter. The precision of the C<sup>14</sup> assay

ranges from 0.5 to 1.0 per cent, except in some of the early measurements in which the errors ranged up to 1.5 per cent. The ratio of stable C<sup>13</sup> to C<sup>12</sup> in the purified CO<sub>2</sub> gas is determined with a precision of 0.1 per cent using a model 201 Consolidated Electrodynamics double-collecting mass spectrometer. The details of this procedure have been published separately [Broecker, Tucek, and Olson, 1959].

The results (Table 1) are ultimately expressed as per mil differences from a standard; the exact expression is

$$\delta C^{14} = \frac{A_{\text{sample}}^* - 0.950 A_{\text{standard}}^0}{0.950 A_{\text{standard}}^0} \times 1000$$

where  $A_{\text{sample}}^*$  is the net C<sup>14</sup> activity of the sample corrected for radioactive decay between the time of collection (or formation, in the case of tree ring samples) and measurement and  $A_{\text{standard}}^0$  is the net C<sup>14</sup> activity of the National Bureau of Standards oxalic acid standard corrected for radioactive decay between January 1, 1958, and the time of measurement. The factor 0.950 is a defined constant introduced to place the zero on the  $\delta C^{14}$  scale close to the value of 19th century wood samples. In order to eliminate differences introduced by fractionation, the results are normalized to a common C<sup>13</sup>/C<sup>12</sup> ratio. The normalization equation is

$$\Delta C^{14} = \delta C^{14} - 2 \delta C^{13} (1 + (\delta C^{14}/1000)) - 50.0$$

where  $\delta C^{14}$  is defined as above and

$$\delta C^{13} = \frac{R_{\text{sample}} - R_{\text{belemnite standard}}}{R_{\text{belemnite standard}}} \times 1000$$

where  $R$  stands for C<sup>13</sup>/C<sup>12</sup> ratio. The constant factor (−50.0) is introduced in order to place the zero on the scale at a convenient point (the C<sup>14</sup>/C<sup>12</sup> ratio corrected in age to pre-1900 wood samples).

The normalization of the results to a common  $\delta C^{13}$  value is of considerable convenience. Not only does it remove differences introduced by fractionation during chemical processing but it also allows results from various materials (i.e., atmospheric CO<sub>2</sub>, ocean HCO<sub>3</sub><sup>−</sup>, and organic materials) to be directly compared as if they had formed within a system free of isotope



fractionation effects. (See Broecker and Olson [1959] for a more complete discussion of the derivation of the above scale and of its use.)

The  $\delta C^{13}$  results show a considerable scatter around the rather uniform value of about  $-2$  per mil found for ocean water by several other workers. Much of this scatter is the result of poor gas dispersion with consequent lowering of the efficiency of  $CO_2$  recovery in some of the KOH bubblers used in shipboard processing. As shown by Craig [1953], during uptake by basic solutions the lighter isotopes are more efficiently captured by the solution. Of the samples reported in this paper those showing poor chemical yield are the ones for which large negative  $\delta C^{13}$  values were observed. In general a rather consistent relationship between  $\delta C^{13}$  and yield is apparent. For 30 of the 135 analyses reported, no  $\delta C^{13}$  values are available. In these cases values have been estimated from yield data and from  $\delta C^{13}$  results on other samples from the same cruise. The errors on the  $\Delta C^{14}$  values computed have been increased in order to cover the uncertainty in these estimates. Since the  $C^{14}$  fractionation will always be twice the  $C^{13}$  fractionation, no significant error is introduced in cases where mass-spectrometer measurements are available.

**Results.** Since 1955, radiocarbon measurements have been made on over 130 sea water samples collected by Lamont workers aboard the research vessel *Vema* in the North and South Atlantic oceans. The results of these measurements are given in Table 1.

On the basis of the findings of numerous oceanographic surveys over the past 35 years, beginning with the *Meteor* expedition, the Atlantic has become the best known of the world's oceans. The broad features of the distribution of water masses and the general circulation of the North and South Atlantic have been summarized by Sverdrup, Johnson, and Fleming [1942] and Sverdrup [1954]. These features, based on the observed distribution of conservative properties, allow the vertical and areal extent of the different water masses, their surface sources, and some of their mixing characteristics to be defined.

The major water masses described in these studies of the Atlantic may be listed as follows:

*North Atlantic Central Water (NACW).* This

water mass lies beneath the surface mixed layer in the central North Atlantic between the depths of 100 and 700 meters. Characterized by relatively high temperature and salinity, it lies within the great gyre of surface circulation bounded on the west by the Gulf Stream, to the north by the North Atlantic Current, to the east by the Canary Current, and on the south by the North Equatorial Current. The boundaries of this mass lie between temperatures of  $8^\circ C$  and  $19^\circ C$  and salinities of 35.1 per mil to 36.7 per mil. Unlike the deeper masses, rather steep temperature and salinity gradients exist between the top and bottom of the mass, suggesting an origin through the mixing of a warm and a cold water mass of different salinity. Whether this mixing occurs at the source or at depth is not known; the resultant gradient can be explained equally well by either process.

*South Atlantic Central Water (SACW).* This water mass is the southern-hemisphere counterpart of the NACW, although somewhat shallower in vertical extent, lying between 100 and 600 meters. The water has a vertical decrease in temperature ranging from  $6^\circ$  to  $18^\circ C$  and salinity from 34.5 to 36.0 per mil. This water outcrops at the surface in a zone known as the Subtropical Convergence, extending along latitude  $40^\circ S$ , and can be traced at depth throughout the South Atlantic to the equator.

*Antarctic Intermediate Water (AAIW).* Water of this description is thought to form along the conspicuous circumpolar belt (in the southern ocean) known as the Antarctic Convergence. In the South Atlantic this convergence lies at latitude  $50^\circ S$ . Here water of  $\sim 3^\circ C$  temperature and salinity of  $\sim 34$  per mil sinks and moves northward at depths centered at 700 to 1000 meters. This water can easily be identified by its origin, having a lower salinity than surrounding waters. As it moves northward it mixes with waters above and below, gradually attaining higher temperature and salinity. Traces of AAIW can be observed in the western North Atlantic as far north as  $30^\circ N$ .

*Antarctic Bottom Water (AABW).* Water formed in the Weddell Sea area of the Antarctic has the highest density of all ocean waters. It can be traced in the western basin of the Atlantic beginning at a temperature of  $0^\circ C$  and a salinity of 36.65 per mil in

| Date Collected                       | Latitude | Longitude | Sample Depth, meters | Bottom Depth, meters | Temp., °C | Salinity, ‰ | Water Type†              | $\delta^{14}\text{C}$ , ‰ | $\delta^{13}\text{C}$ , ‰ | $\Delta^{14}\text{C}$ , ‰ | Index No. |
|--------------------------------------|----------|-----------|----------------------|----------------------|-----------|-------------|--------------------------|---------------------------|---------------------------|---------------------------|-----------|
| <b>I. Surface Water, 0-50 meters</b> |          |           |                      |                      |           |             |                          |                           |                           |                           |           |
| <b>A. North Atlantic</b>             |          |           |                      |                      |           |             |                          |                           |                           |                           |           |
| 10/30/55                             | 38°22'N  | 71°32'W   | 0                    | 2910                 | 17.9      | 36.35       | Gulf Stream              | -9 ± 7                    | +1.8                      | -62 ± 7                   | 1         |
| 7/19/55                              | 37°58'N  | 50°53'W   | 0                    | 5423                 | 24.9      |             | North Atlantic Current   |                           |                           |                           |           |
| 11/ 1/55                             | 36°06'N  | 66°06'W   | 0                    | 5158                 | 23.1      |             | Sargasso Sea             | +12 ± 8                   | (-2.0)                    | (-34 ± 10)                | 2         |
| 11/ 3/55                             | 34°06'N  | 65°06'W   | 0                    | 5069                 | 24.8      |             | Sargasso Sea             | -5 ± 6                    | -1.1                      | -52 ± 7                   | 3         |
| 11/13/57                             | 34°05'N  | 65°00'W   | 0                    | 4726                 | 22.2      |             | Sargasso Sea             | +4 ± 6                    | (+5.0)                    | (-56 ± 8)                 | 4         |
| 6/ 2/56                              | 33°50'N  | 66°18'W   | 0                    | 5114                 | 21.7      |             | Sargasso Sea             | -1 ± 5                    | +1.5                      | -54 ± 5                   | 5         |
| 6/ 7/56                              | 33°00'N  | 49°48'W   | 0                    | 4667                 | 22.3      |             | Sargasso Sea             | -5 ± 5                    | +2.2                      | -59 ± 5                   | 6         |
| 6/ 5/56                              | 32°38'N  | 57°55'W   | 0                    | 5218                 | 22.7      |             | Sargasso Sea             | -6 ± 6                    | +0.7                      | -58 ± 6                   | 7         |
| 6/12/56                              | 31°45'N  | 34°38'W   | 0                    | 3878                 | 21.9      |             | Sargasso Sea             | +9 ± 6                    | +1.4                      | -44 ± 6                   | 8         |
| 11/16/57                             | 29°57'N  | 61°41'W   | 0                    | 2748                 | 23.3      |             | Sargasso Sea             | -13 ± 7                   | (+1.0)                    | (-65 ± 10)                | 9         |
| 11/17/57                             | 29°13'N  | 60°30'W   | 46                   | 5836                 | 24.2      | 36.56       | Sargasso Sea             | +26 ± 5                   | (+5.0)                    | (-34 ± 8)                 | 10        |
| 11/ 5/55                             | 27°05'N  | 73°32'W   | 0                    | 4717                 | 25.1      |             | Sargasso Sea             | +13 ± 6                   | -1.8                      | -33 ± 6                   | 11        |
| 6/24/57                              | 25°42'N  | 79°23'W   | 0                    | ...                  | ...       |             | Antilles Current         | +5 ± 6                    | +4.9                      | -55 ± 6                   | 12        |
| 11/ 8/55                             | 25°25'N  | 75°13'W   | 0                    | 4561                 | 26.1      |             | Florida Current          | -2 ± 7                    | +2.3                      | -44 ± 7                   | 13        |
| 11/12/55                             | 20°32'N  | 68°30'W   | 0                    | 5091                 | 27.8      |             | Antilles Current         | +6 ± 5                    | +0.8                      | -45 ± 5                   | 14        |
| 11/10/55                             | 19°58'N  | 70°53'W   | 0                    | 1801                 | 28.2      |             | Antilles Current         | +10 ± 10                  | +2.0                      | -44 ± 10                  | 15        |
| 11/14/55                             | 19°07'N  | 67°07'W   | 0                    | 4973                 | 27.8      |             | Antilles Current         | +14 ± 7                   | +7.5                      | -51 ± 7                   | 16        |
| 11/26/57                             | 15°03'N  | 39°48'W   | 0                    | 5222                 | 26.7      |             | North Equatorial Current | +8 ± 5                    | +2.1                      | -46 ± 5                   | 17        |
| <b>B. Caribbean</b>                  |          |           |                      |                      |           |             |                          |                           |                           |                           |           |
| 12/20/55                             | 19°04'N  | 80°48'W   | 0                    |                      | 27.3      |             | West Caribbean           | +1 ± 8                    | -1.0                      | -47 ± 8                   | 18        |
| 12/10/55                             | 16°17'N  | 79°14'W   | 0                    | 1324                 | 27.6      |             | West Caribbean           | -1 ± 7                    | (+3.0)                    | (-57 ± 9)                 | 19        |
| 12/16/55                             | 17°39'N  | 79°04'W   | 0                    | 798                  | ...       |             | West Caribbean           | -11 ± 10                  | +7.2                      | -76 ± 10                  | 20        |
| 12/ 8/55                             | 12°27'N  | 77°25'W   | 0                    | 3918                 | 27.8      |             | East Caribbean           | -5 ± 7                    | -1.1                      | -53 ± 7                   | 21        |
| 11/25/55                             | 17°05'N  | 71°36'W   | 0                    | 2075                 | 27.8      |             | East Caribbean           | +9 ± 6                    | +2.3                      | -46 ± 6                   | 22        |
| 11/19/55                             | 16°43'N  | 70°38'W   | 0                    | 4101                 | 27.8      |             | East Caribbean           | +9 ± 7                    | +5.5                      | -52 ± 7                   | 23        |
| 11/17/55                             | 17°11'N  | 68°55'W   | 0                    | 5567                 | 27.4      |             | East Caribbean           | +4 ± 7                    | +4.6                      | -55 ± 7                   | 24        |
| 11/22/55                             | 17°47'N  | 68°22'W   | 0                    | 2536                 | 27.8      |             | East Caribbean           | -1 ± 6                    | +5.1                      | -61 ± 6                   | 25        |
| <b>C. Mediterranean</b>              |          |           |                      |                      |           |             |                          |                           |                           |                           |           |
| 8/16/56                              | 34°04'N  | 26°21'E   | 0                    | 3181                 | 26.0      |             | E. Mediterranean         | +11 ± 7                   | +5.9                      | -51 ± 7                   | 26        |
| 7/24/56                              | 38°06'N  | 22°54'E   | 0                    | 795                  | 25.2      |             | Gulf of Corinth          | -4 ± 7                    | +4.2                      | -63 ± 7                   | 27        |
|                                      |          |           |                      |                      |           |             |                          | -53 ± 7                   | -2.3                      | -99 ± 7                   | 28        |

TABLE 1. Continued

| Date Collected    | Latitude | Longitude | Sample Depth, meters | Bottom Depth, meters | Temp., °C | Salinity, ‰ | Water Type†              | $\delta C^{14}$ , ‰ | $\delta C^{13}$ , ‰ | $\Delta C^{14}$ , ‰ | Index No. |
|-------------------|----------|-----------|----------------------|----------------------|-----------|-------------|--------------------------|---------------------|---------------------|---------------------|-----------|
| D. South Atlantic |          |           |                      |                      |           |             |                          |                     |                     |                     |           |
| 12/ 5/57          | 00°51'N  | 32°52'W   | 49                   | 3303                 | 26.9      | 36.08       | South Equatorial Current | -28 ± 6             | -1.1                | -76 ± 6             | 29        |
| 2/15/56           | 03°06'S  | 32°26'W   | 0                    | 4228                 | 27.6      |             | South Equatorial Current | -2 ± 6              | +1.4                | -55 ± 6             | 30        |
| 2/17/56           | 03°34'S  | 31°22'W   | 0                    | 4411                 | 26.9      |             | South Equatorial Current | -13 ± 7             | (+1.0)              | (-61 ± 10)          | 31        |
| 1/ 9/57           | 04°29'S  | 34°54'W   | 0                    | 3371                 | 27.1      |             | South Equatorial Current | -24 ± 5             | -0.4                | -73 ± 5             | 32        |
| 5/20/57           | 05°41'S  | 10°39'E   | 0                    | 2087                 | 27.5      |             | North Benguela Current   | -24 ± 7             | -3.3                | -68 ± 7             | 33        |
| 12/16/57          | 09°38'S  | 34°05'W   | 0                    | 4707                 | 27.2      |             | Brazil Current           | -8 ± 7              | (-1.0)              | (-56 ± 9)           | 34        |
| 1/19/57           | 10°59'S  | 32°28'W   | 0                    | 4978                 | 27.2      |             | Brazil Current           | -11 ± 7             | -4.2                | -52 ± 7             | 35        |
| 12/26/57          | 23°12'S  | 37°38'W   | 0                    | 3824                 | 26.4      |             | Brazil Current           | +7 ± 7              | -2.1                | -39 ± 7             | 36        |
| 5/12/57           | 14°30'S  | 07°34'E   | 0                    | 3985                 | 24.4      |             | Central South Atlantic   | -10 ± 15            | (-1.0)              | (-58 ± 15)          | 37        |
| 5/ 5/57           | 22°34'S  | 14°12'E   | 73                   | Plankton sample      |           |             | Central South Atlantic   | -64 ± 15            | -16.0               | -84 ± 15            | 38        |
| 5/ 3/57           | 25°31'S  | 12°26'E   | 0                    | 3030                 | 17.9      |             | Central South Atlantic   | -13 ± 5             | -1.3                | -60 ± 5             | 39        |
| 4/29/57           | 32°13'S  | 16°20'E   | 0                    | 455                  | 19.3      |             | Central South Atlantic   | -2 ± 7              | -0.9                | -51 ± 7             | 40        |
| 4/ 6/58           | 34°06'S  | 18°06'E   | 0                    | 188                  | 20.0      |             | Central South Atlantic   | -8 ± 6              | -1.2                | -56 ± 6             | 41        |
| 4/19/57           | 34°46'S  | 06°29'E   | 0                    | 5308                 | 17.0      |             | West Benguela            | -2 ± 7              | -5.1                | -42 ± 7             | 42        |
| 4/ 6/57           | 39°03'S  | 41°48'W   | 0                    | 5112                 | 17.8      |             | West Wind Drift          | -2 ± 5              | +1.0                | -54 ± 5             | 43        |
| 4/12/57           | 40°54'S  | 20°29'W   | 0                    | 3351                 | 12.3      |             | West Wind Drift          | -21 ± 7             | -6.6                | -58 ± 7             | 44        |
| 4/15/57           | 41°15'S  | 06°10'W   | 0                    | 4085                 | 12.3      |             | West Wind Drift          | -14 ± 7             | -2.2                | -59 ± 7             | 45        |
| 4/ 2/57           | 40°43'S  | 56°32'W   | 0                    | 1022                 | 11.1      |             | Falkland Current         | -27 ± 5             | +0.6                | -79 ± 5             | 46        |
| 4/ 3/57           | 41°05'S  | 51°09'W   | 0                    | 5506                 | 16.6      |             | Falkland Current         | -27 ± 7             | -1.6                | -74 ± 7             | 47        |
| 3/20/57           | 45°24'S  | 59°13'W   | 0                    | 1726                 | 9.0       |             | Falkland Current         | -27 ± 6             | +0.3                | -78 ± 6             | 48        |
| E. Antarctic      |          |           |                      |                      |           |             |                          |                     |                     |                     |           |
| 3/24/58           | 51°27'S  | 02°38'E   | 0                    | 3329                 | 3.5       |             | Antarctic                |                     |                     |                     |           |

| Date Collected                   | Latitude | Longitude | Depth, meters  | Depth, meters | Temp., °C | Salinity, ‰ | Water Type†             | $\delta C^{13}$ , ‰ | $\delta C^{14}$ , ‰ | Index No. |
|----------------------------------|----------|-----------|----------------|---------------|-----------|-------------|-------------------------|---------------------|---------------------|-----------|
| 2/24/58                          | 55°27'S  | 57°10'W   | 0              | 1750          | 3.0       |             | Drake Passage           | -69 ± 7             | -111 ± 7            | 50        |
| 3/19/58                          | 57°07'S  | 07°15'W   | 0              | 3444          | 1.0       |             | Antarctic Surface Water | -79 ± 5             | -120 ± 5            | 51        |
| II. Subsurface Water, below 50 m |          |           |                |               |           |             |                         |                     |                     |           |
| A. 200-400 meters                |          |           |                |               |           |             |                         |                     |                     |           |
| 4/16/56                          | 25°01'N  | 59°12'W   | 256            | 6081          | (18)      | (36.6)      | NACW                    | -11 ± 7             | +3.1                | 52        |
| 10/12/56                         | 23°22'N  | 27°14'W   | 329            | 5499          | (16)      | 36.08       | NACW                    | -15 ± 5             | +4.2                | 53        |
| 4/11/56                          | 15°49'N  | 47°12'W   | 229            | 3757          | (16)      | (36.2)      | NACW                    | -29 ± 6             | -2.6                | 54        |
| 4/7/56                           | 08°15'N  | 37°54'W   | 274            | 3520          | (9)       | (34.9)      | Transition              |                     |                     |           |
| 12/3/57                          | 07°22'N  | 29°59'W   | 240            | 4894          | (9)       | 35.10       | NACW-SACW Transition    | -49 ± 9             | -2.9                | 55        |
| 6/9/57                           | 00°19'S  | 24°29'W   | 201            | 2988          | 13.5      | (35.2)      | NACW-SACW               | -41 ± 6             | -0.9                | 56        |
| 2/16/56                          | 03°17'S  | 32°14'W   | 274            | 4483          | (10)      | 34.91       | SACW                    | -18 ± 7             | -2.4                | 57        |
| 5/19/57                          | 06°25'S  | 11°26'W   | 366            | 439           | 10.1      | (35.0)      | SACW                    | -35 ± 7             | -0.4                | 58        |
| 12/21/57                         | 14°02'S  | 37°32'W   | 295            | 3804          | 16.35     | 35.71       | Lower SACW              | -38 ± 7             | -2.5                | 59        |
| 12/21/57                         | 14°10'S  | 37°33'W   | 300            | 3749          | (15)      | 35.37       | SACW                    | -8 ± 5              | -1.9                | 60        |
| 5/9/57                           | 21°14'S  | 03°22'E   | Surface to 274 | 5134          | (11)      | (34.9)      | SACW                    | -4 ± 6              | +1.6                | 61        |
|                                  |          |           |                |               |           |             |                         | -13 ± 7             | -1.5                | 62        |
| 5/2/57                           | 28°26'S  | 08°39'E   | 201            | 5066          | (14)      | (35.3)      | SACW                    | -4 ± 7              | -2.5                | 63        |
| B. 500-1200 meters               |          |           |                |               |           |             |                         |                     |                     |           |
| 7/19/55                          | 38°00'N  | 51°37'W   | 732            | 5405          | (12)      | 35.95       | Lower NACW              | +7 ± 13             | (-2.0)              | 64        |
| 10/11/56                         | 24°23'N  | 24°03'W   | 1097           | 5104          | (7)       | 35.18       | N. Atlantic             | -41 ± 10            | +8.5                | 65        |
| 6/17/55                          | 23°28'N  | 65°56'W   | 896            |               | (8)       | 35.12       | Undefined N. Atlantic   | -50 ± 13            | (-2.0)              | 66        |
| 4/14/56                          | 22°22'N  | 54°19'W   | 823            | 5654          | (8)       | (35.0)      | Undefined N. Atlantic   | -49 ± 6             | (-2.0)              | 67        |
| 11/22/57                         | 20°44'N  | 49°24'W   | 950            | 4354          | 6.0       | 34.97       | Undefined N. Atlantic   | -57 ± 6             | -14.8               | 68        |
| 12/25/55                         | 19°09'N  | 76°59'W   | 1097           | 3749          | (5)       | 34.84       | Undefined N. Atlantic   | -43 ± 7             | -1.3                | 69        |
| 4/11/56                          | 15°49'N  | 47°12'W   | 823            | 3757          | (6)       | (34.7)      | Undefined N. Atlantic   | -50 ± 6             | 0.0                 | 70        |
| 11/26/57                         | 15°29'N  | 40°30'W   | 823            | 4330          | 6.1       | 34.85       | Undefined N. Atlantic   | -73 ± 6             | -7.1                | 71        |



TABLE 1. Continued

| Date Collected      | Latitude | Longitude | Sample Depth, meters | Bottom Depth, meters | Temp., °C | Salinity, ‰ | Water Type†           | $\delta C^{14}$ , ‰ | $\delta C^{13}$ , ‰ | $\Delta C^{14}$ , ‰ | Index No. |
|---------------------|----------|-----------|----------------------|----------------------|-----------|-------------|-----------------------|---------------------|---------------------|---------------------|-----------|
| 4/ 7/56             | 08°14'N  | 37°48'W   | 823                  | 3598                 | (5)       | (34.7)      | N. Atlantic Undefined | -60 ± 5             | -0.1                | -110 ± 5            | 72        |
| 12/ 5/57            | 00°51'N  | 32°52'W   | 799                  | 3314                 | (4.3)     | 34.60       | AAIW                  | -63 ± 6             | -1.3                | -111 ± 6            | 73        |
| 2/16/56             | 03°22'S  | 31°50'W   | 732                  | 4301                 | (4.4)     | 34.46       | AAIW                  | -76 ± 6             | -0.2                | -125 ± 6            | 74        |
| 5/26/57             | 04°47'S  | 02°40'E   | 732                  | 5238                 | (5)       | (34.5)      | AAIW                  | -64 ± 7             | -12.7               | -90 ± 7             | 75        |
| 12/16/57            | 10°04'S  | 33°52'W   | 914                  | 4625                 | (4.0)     | 34.65       | AAIW                  | -86 ± 6             | -0.7                | -134 ± 6            | 76        |
| 5/ 9/57             | 21°09'S  | 03°20'E   | 550                  | 5084                 | 5.96      | (34.5)      | Upper AAIW            | -67 ± 7             | -2.1                | -113 ± 9            | 77        |
| 5/ 2/57             | 28°25'S  | 08°36'E   | 1189                 | 5060                 | 3.5       | (34.5)      | Lower AAIW            | -92 ± 7             | -2.4                | -137 ± 7            | 78        |
| C. 1200-2500 meters |          |           |                      |                      |           |             |                       |                     |                     |                     |           |
| 10/31/57            | 39°16'N  | 70°46'W   | 2323                 | 2642                 | 3.2       | 34.88       | Upper NADW            | -22 ± 5             | -2.2                | -68 ± 5             | 80        |
| 7/19/55             | 37°58'N  | 50°53'W   | 1944                 | 5423                 | 3.69      | 34.88       | Upper NADW            | -19 ± 7             | (-3.0)              | (-65 ± 13)          | 81        |
| 11/17/57            | 29°13'N  | 60°31'W   | 1423                 | 5843                 | 4.75      | 35.25       | Upper NADW            | -32 ± 10            | (-2.0)              | (-76 ± 10)          | 82        |
| 6/15/55             | 25°07'N  | 69°57'W   | 1829                 | 5504                 | (3.8)     | 34.94       | Upper NADW            | -20 ± 13            | -1.7                | -67 ± 13            | 83        |
| 11/ 4/57            | 20°31'N  | 73°22'W   | 1230                 | 3260                 | (4)       | (35.0)      | Upper NADW            | -5 ± 6              | -2.7                | -49 ± 6             | 84        |
| 7/ 4/57             | 11°12'N  | 59°19'W   | 2305                 | 4113                 | (3)       | (35.0)      | Upper NADW            | -61 ± 7             | -13.0               | -88 ± 7             | 87        |
| 7/ 4/57             | 11°12'N  | 57°16'W   | 1765                 | 4114                 | (4)       | (35.0)      | Upper NADW            | -43 ± 6             | -5.6                | -83 ± 6             | 88        |
| 7/ 3/57             | 10°49'N  | 55°41'W   | 1262                 | 4396                 | (5)       | (34.9)      | Upper NADW            | -62 ± 6             | -15.9               | -83 ± 6             | 89        |
| 10/10/56            | 25°13'N  | 21°23'W   | 1463                 | 4548                 | (5)       | 35.07       | Upper NADW            | -56 ± 5             | -0.7                | -104 ± 5            | 90        |
| 9/10/56             | 25°13'N  | 21°23'W   | 1829                 | 4656                 | (4.5)     | 35.08       | Upper NADW            | -60 ± 6             | +0.3                | -111 ± 6            | 91        |
| 4/ 7/56             | 08°15'N  | 37°53'W   | 1463                 | 3598                 | (4)       | (35.0)      | Upper NADW            | -53 ± 7             | +2.5                | -108 ± 7            | 92        |
| 5/27/57             | 04°23'S  | 00°05'W   | 1829                 | 4799                 | 3.8       | (34.96)     | Upper NADW            | -47 ± 6             | -0.8                | -96 ± 6             | 93        |
| 12/16/57            | 10°06'S  | 34°54'W   | 1829                 | 4625                 | 3.83      | 34.96       | Upper NADW            | -54 ± 8             | -0.9                | -102 ± 8            | 94        |
| 12/17/57            | 12°37'S  | 35°00'W   | 1829                 | 4378                 | (3.7)     | 34.96       | Upper NADW            | -55 ± 9             | -6.6                | -92 ± 9             | 95        |
| 12/22/57            | 15°17'S  | 36°00'W   | 1900                 | 4369                 | 3.6       | 34.94       | Upper NADW            | -61 ± 7             | +0.4                | -112 ± 7            | 96        |
| 5/ 8/57             | 22°40'S  | 03°16'E   | 1646                 | 5046                 | (3.2)     | (34.8)      | Antarctic and NADW    | -80 ± 10            | -3.3                | -124 ± 10           | 97        |
| 4/ 2/57             | 35°20'S  | 10°42'E   | 2195                 | 5008                 | (2.8)     | (34.8)      | Antarctic and NADW    | -71 ± 10            | +0.7                | -122 ± 10           | 98        |
| 4/ 7/57             | 38°58'S  | 40°06'W   | 1829                 | 5103                 | 2.98      | (34.7)      | Antarctic and NADW    | -100 ± 5            | -4.3                | -143 ± 5            | 99        |
| D. 2500-4000 meters |          |           |                      |                      |           |             |                       |                     |                     |                     |           |
| 7/27/55             | 34°55'N  | 57°11'W   | 3488                 | 3671                 | 2.70      | 34.91       | NADW                  | -43 ± 10            | (-2.0)              | (-89 ± 10)          | 100       |
| 4/18/56             | 28°41'N  | 60°19'W   | 2560                 | 5528                 | (3.3)     | (34.95)     | NADW                  | -62 ± 6             | +0.9                | -114 ± 6            | 101       |
| 9/ 9/56             | 26°10'N  | 18°06'W   | 3503                 | 3548                 | (2.6)     | 34.91       | NADW                  | -62 ± 8             | +0.6                | -113 ± 8            | 102       |
| 7/ 9/56             | 26°10'N  | 18°06'W   | 3503                 | 3548                 | (2.6)     | 34.91       | NADW                  | -62 ± 8             | +0.6                | -113 ± 8            | 103       |

| Collected        | Lat-<br>tude | Longi-<br>tude | Depth,<br>meters | Depth,<br>meters | Temp.,<br>°C | Salinity,<br>‰ | Water Type†     | $\delta C^{13}$ ,<br>‰ | $\delta C^{14}$ ,<br>‰ | Index<br>No. |
|------------------|--------------|----------------|------------------|------------------|--------------|----------------|-----------------|------------------------|------------------------|--------------|
| 11/21/57         | 22°03'N      | 51°27'W        | 2700             | 5050             | 2.98         | 34.96          | NADW            | -53 ± 6                | -5.3                   | 105          |
| 11/22/57         | 20°43'N      | 49°26'W        | 2926             | 4058             | 2.92         | 34.96          | NADW            | -62 ± 6                | -1.3                   | 106          |
| 11/25/57         | 16°44'N      | 42°38'W        | 3840             | 4502             | 2.46         | 34.92          | NADW            | -69 ± 7                | -3.9                   | 107          |
| 12/31/56         | 15°52'N      | 38°03'W        | 3840             | 5300             | (2.5)        | 34.95          | NADW            | -53 ± 5                | -9.6                   | 108          |
| 4/11/56          | 15°49'N      | 47°12'W        | 2560             | 3757             | (3.0)        | (34.95)        | NADW            | -74 ± 7                | +3.1                   | 109          |
| 2/18/55          | 03°48'S      | 30°53'W        | 2560             | 4426             | (2.9)        | (34.95)        | NADW            | -49 ± 6                | -1.0                   | 110          |
| 5/27/57          | 04°23'S      | 00°15'W        | 2743             | 4779             | 2.8          | (34.89)        | NADW            | -62 ± 7                | -1.3                   | 111          |
| 1/22/57          | 17°03'S      | 28°13'W        | 2807             | 5141             | (2.75)       | 34.93          | NADW            | -71 ± 6                | -8.3                   | 112          |
| 5/8/57           | 22°56'S      | 04°52'E        | 2560             | 2824             | (2.75)       | (34.88)        | NADW            | -63 ± 7                | -3.8                   | 113          |
| 1/18/58          | 38°25'S      | 53°11'W        | 2770             | 2908             | 2.75         | 34.89          | NADW            | -57 ± 6                | -2.3                   | 114          |
| E. > 4000 meters |              |                |                  |                  |              |                |                 |                        |                        |              |
| 7/16/55          | 39°27'N      | 56°57'W        | 4345             | 5260             | 2.32         | 34.80          | NADW (W. Basin) | -43 ± 7                | (-2.0)                 | 115          |
| 7/18/55          | 39°03'N      | 53°29'W        | 5281             | 5317             | 2.26         | (34.9)         | NADW (W. Basin) | -58 ± 7                | (-2.0)                 | 116          |
| 7/19/55          | 38°00'N      | 51°37'W        | 5369             | 5405             | 2.31         | 34.80          | NADW (W. Basin) | -40 ± 13               | (-86 ± 13)             | 117          |
| 7/23/55          | 37°23'N      | 53°22'W        | 5368             | 5404             | 2.25         | 34.80          | NADW (W. Basin) | -73 ± 13               | (-119 ± 13)            | 118          |
| 7/26/55          | 35°28'N      | 55°48'W        | 5416             | 5452             | 2.60         | 34.80          | NADW (W. Basin) | -53 ± 7                | (-99 ± 9)              | 119          |
| 7/22/55          | 35°43'N      | 53°15'W        | 5454             | 5490             | 2.25         | 34.80          | NADW (W. Basin) | -58 ± 5                | (-104 ± 9)             | 120          |
| 7/21/55          | 34°46'N      | 52°46'W        | 5481             | 5517             | (2.2)        | 34.81          | NADW-(AABW)     | -76 ± 13               | (-2.0)                 | 121          |
| 6/27/55          | 31°47'N      | 71°13'W        | 5360             | 5405             | 2.16         | 34.80          | NADW-(AABW)     | -60 ± 13               | (-106 ± 13)            | 122          |
| 6/29/55          | 31°21'N      | 66°39'W        | 4893             | 4938             | 2.22         | 34.86          | NADW-(AABW)     | -56 ± 7                | (-102 ± 9)             | 123          |
| 6/27/55          | 31°17'N      | 71°03'W        | 5367             | 5412             | 2.19         | 35.17          | NADW-(AABW)     | -52 ± 9                | (-98 ± 10)             | 124          |
| 6/26/55          | 29°14'N      | 69°55'W        | 5400             | 5447             | 2.11         | (34.9)         | NADW-(AABW)     | -33 ± 13               | (-79 ± 13)             | 125          |
| 6/15/55          | 25°07'N      | 69°57'W        | 5508             | 5564             | (2.2)        | 34.74          | NADW-(AABW)     | -53 ± 20               | (-99 ± 20)             | 127          |
| 4/16/56          | 25°01'N      | 59°12'W        | 6035             | 6081             | (2.0)        | (34.9)         | NADW-(AABW)     | -88 ± 6                | -7.3                   | 128          |
| 10/11/56         | 24°23'N      | 24°03'W        | 4921             | 5104             | (2.5)        | 34.92          | NADW (E. Basin) | -82 ± 5                | +1.5                   | 129          |
| 6/20/55          | 24°05'N      | 68°23'W        | 5584             | 5649             | (2.2)        | (34.9)         | NADW-(AABW)     | -58 ± 8                | (-2.0)                 | 130          |
| 6/17/55          | 23°28'N      | 65°56'W        | 5788             | 5835             | (2.5)        | 34.76          | NADW-(AABW)     | -55 ± 13               | (-101 ± 13)            | 131          |
| 4/14/56          | 22°22'N      | 54°19'W        | 5608             | 5654             | (2.0)        | (34.9)         | NADW-(AABW)     | -87 ± 6                | -6.6                   | 132          |
| 12/31/56         | 15°52'N      | 38°08'W        | 5264             | 5300             | (2.4)        | 34.86          | NADW (E. Basin) | -65 ± 7                | +2.3                   | 133          |
| 11/26/57         | 15°29'N      | 40°30'W        | 4147             | 4330             | 2.45         | 34.67          | NADW (E. Basin) | -83 ± 9                | -126 ± 9               | 134          |
| 2/16/56          | 03°26'S      | 31°34'W        | 4389             | 4429             | (0.8)        | (34.80)        | AABW            | -99 ± 7                | +0.6                   | 135          |
| 1/22/57          | 16°34'S      | 28°03'W        | 5330             | 5376             | (0.5)        | 34.69          | AABW            | -105 ± 6               | -2.5                   | 136          |
| 2/1/57           | 33°43'S      | 45°18'W        | 4380             | 4426             | (0.0)        | (34.7)         | AABW            | -88 ± 5                | -3.3                   | 137          |

\* Values listed in parentheses were estimated from published oceanographic data for the location and depth of the sample. Direct measurements are not available.

† See text for definition of abbreviations.

‡  $\delta C^{13}$  values in parentheses are estimated rather than measured. The corresponding  $\Delta C^{14}$  values are based on these estimates.

southern latitudes and gradually becoming warmer and more saline as it mixes with overlying water and moves northward. Traces of this water, highly diluted, can be observed as far as 35°N in the western North Atlantic. Because this water mass lies along the ocean bottom its thickness is highly variable, depending upon the bottom topography. Its upper limits extend to about 3600 meters at 40°S, dropping to 4500 meters at 25°N. Small amounts of AABW are believed to enter the eastern Atlantic basin through deep gaps in the Mid-Atlantic Ridge located at the equator and at 10°N.

*North Atlantic Deep Water (NADW).* This water mass is the most extensive in the Atlantic and is thought to form at the surface in the region of southern Greenland. In this area cool waters from the north mix with Atlantic surface water of higher salinity to form a mixture having a temperature of approximately 2.5°C and a salinity of 34.9 per mil. This water sinks and moves southward to occupy the depths between 1500 and 4000 meters throughout the Atlantic, extending nearly to the Antarctic region.

The samples included in the present survey were collected as nearly as possible from regions where useful correlations with present water-mass concepts could be made. Where possible, independent hydrographic measurements were made to establish a firm relationship between the water sampled with the large-volume sampler and the water masses in the area. In three cases where depth control is considered inadequate, the results have not been used. The measurements presented apply to samples that have a maximum estimated depth error of  $\pm 50$  meters. This depth uncertainty is in no way critical to the conclusions based on these data.

The locations of the surface water samples measured as part of this study are shown in Figure 2. The first number beside each sample location is the chart index number from Table 1 and the second number indicates the  $\Delta C^{14}$  value. Major currents together with directional arrows are shown in a highly generalized fashion.

The most striking feature in the pattern of surface sample results is the systematic change from lower to higher  $C^{14}/C^{12}$  ratios in the samples taken from the Antarctic northward. The effect appears to continue farther north than is

indicated by the samples on this chart, the results of *Fonselius and Ostlund* [1959] for samples from between 60° and 80°N average  $-35$  per mil.

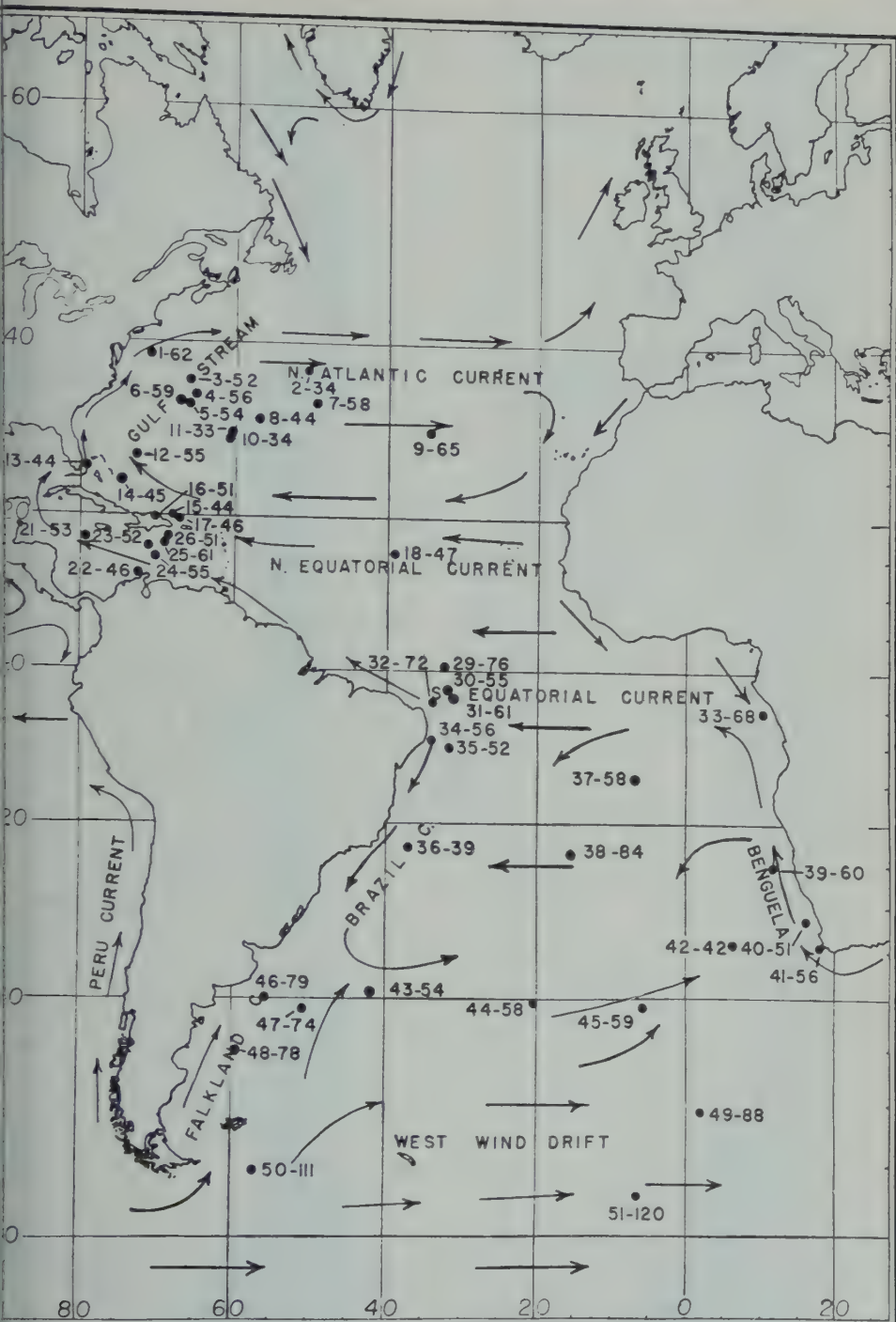
The one sample (No. 51) taken south of the Antarctic Convergence yielded the lowest  $C^{14}$  ratio of all surface samples. This finding is consistent with that of *Burling and Garneault* [press], who report that the lowest surface activities were found in samples from south of the Antarctic Convergence in the region adjacent to New Zealand. The explanation here is the same as that for the South Pacific samples. Deep water which rises to the surface in the circumpolar region south of the Antarctic Convergence is returned to depth before its  $C^{14}$  content can reach equilibrium with the atmosphere.

The average  $\Delta C^{14}$  value for 16 South Atlantic surface samples (excluding the Antarctic and Falkland Current samples) is  $-57$  per mil. Although variations exist they show no systematic pattern. The Falkland Current appears to be water of low  $C^{14}/C^{12}$  ratio northward from Drake Passage. The southward-flowing Brazil Current water has a considerably higher  $C^{14}$  ratio. Where the two meet and move southward as the Brazil Current Extension the  $C^{14}$  content appears to have an intermediate value. It is tempting to assign some significance to the apparent lowering of the surface  $C^{14}$  values from west to east across the Benguela Current region near South Africa. It is known that extensive upwelling of deeper water occurs on the coastal side of the current. Whereas the carbon data are consistent with this circulation pattern the samples are too few and the differences too small to justify any independent conclusions.

The average of 18 samples from the southern North Atlantic is  $-49$  per mil, a difference of 8 per mil from the South Atlantic surface average. The Caribbean has a value (based on ten measurements) intermediate between the two. This is to be expected, since it receives water from both North and South Atlantic mixed layers. Among the samples shown for the North Atlantic there is no apparent correlation of  $\Delta C^{14}$  values with local hydrographic features.

The locations of subsurface samples measured as part of this study are shown in Figure 3.





Locations of surface water samples. The first number beside each dot is the index number and the second the  $\Delta C^4$  value. The arrows show the pattern of surface water movement.



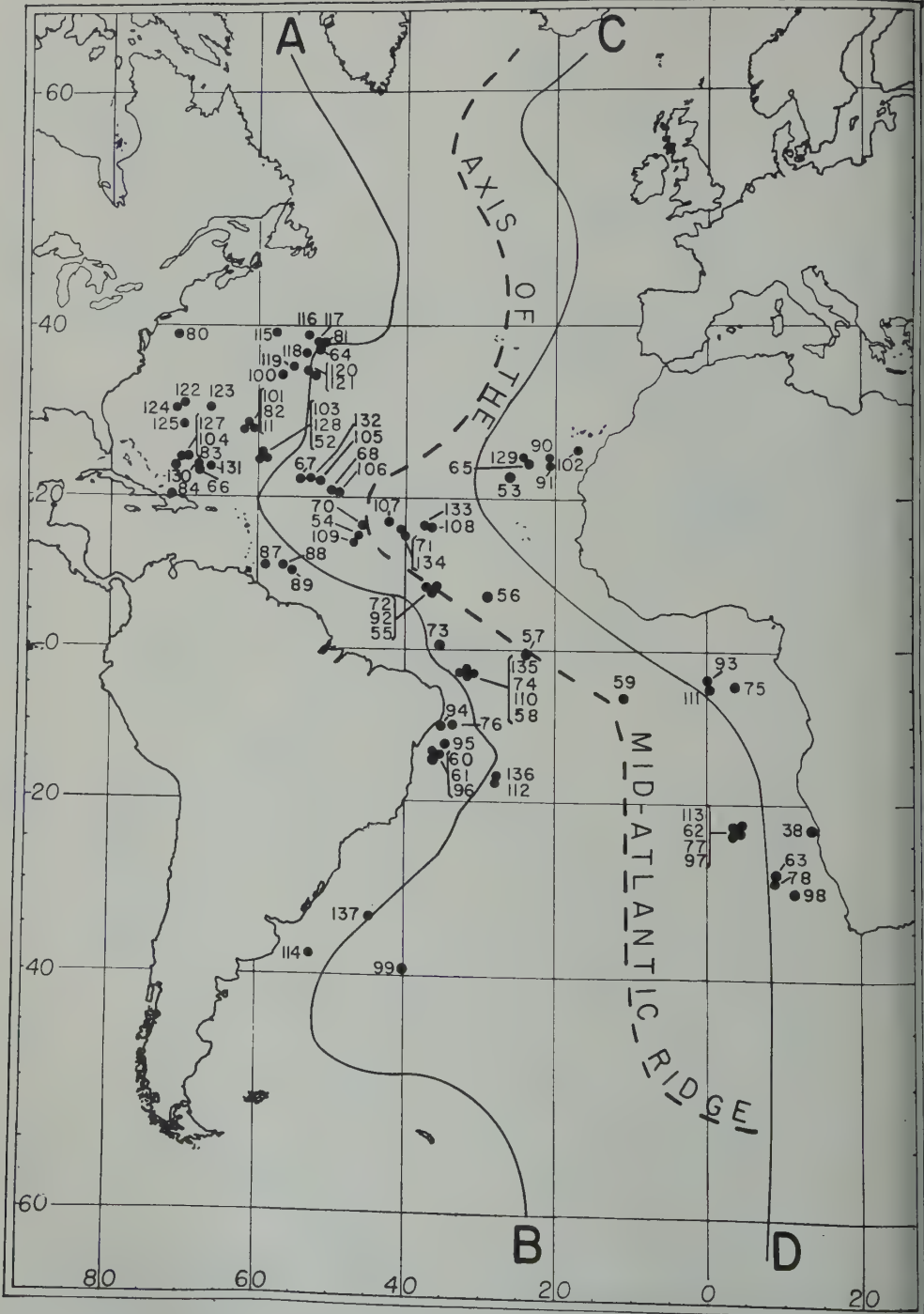


Fig. 3. Locations of subsurface water samples. The number beside each dot is the index number given in Table 1. The lines A-B and C-D give the location of the depth profiles shown in Figure 4.

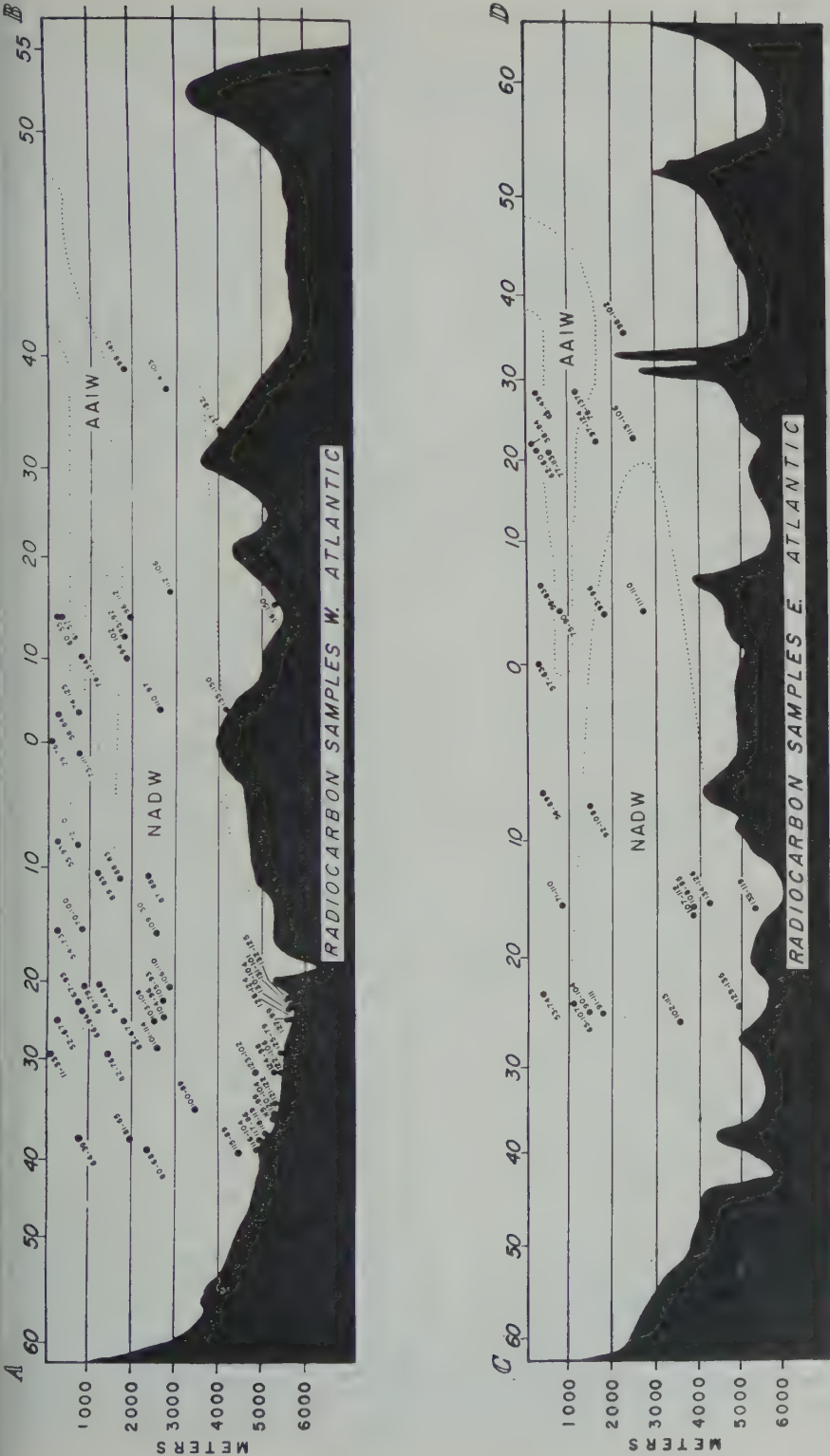


Fig. 4. Locations of subsurface ocean water samples projected onto north-south profiles through the western basin (upper) and eastern basin (lower) of the Atlantic Ocean. The first number beside each dot is the sample index number and the second the  $\Delta C^{14}$  value. The location of the profiles is given in Figure 3.

Their depth distribution is illustrated on two longitudinal cross sections in Figure 4; one drawn through the eastern and one through the western trough of the Atlantic Ocean. Table 5 gives average  $\Delta C^{14}$  values for the results from the various types of water.

In the North Atlantic between  $15^\circ$  and  $30^\circ\text{N}$  and at a depth between 200 and 400 meters three measurements were made, the  $C^{14}$  values averaging  $-71$  per mil. This average, which should be representative of the NACW, is significantly lower than that for the overlying surface water and higher than that for the underlying water.

Eight samples in the North Atlantic from a depth of 800 to 1100 meters average  $-99$  per mil. These samples probably consist of a mixture of AAIW with water of higher temperature and salinity from the North Atlantic and the Mediterranean, the percentage of the latter two increasing northward. Such a distribution is supported by the  $C^{14}$  values which show a northward increase in  $C^{14}/C^{12}$  ratio. At this same level at  $38^\circ\text{N}$  a much higher  $C^{14}/C^{12}$  ratio was obtained (sample 64,  $\Delta C^{14} = -39$  per mil), which suggests not only that the AAIW has ceased to be an influence but also that water which might be defined as Arctic Intermediate Water is present.

In the western Atlantic between  $10^\circ$  and  $40^\circ\text{N}$  and from 1200 to 2400 meters in depth the  $C^{14}$  average for a group of eight samples is  $-72$  per mil. This group appears to represent a  $C^{14}$  maximum level, perhaps being fed by eddy currents emanating from the western boundary current. [Wüst, 1935; Stommel, 1958; Swallow and Worthington, 1957].

In the eastern North Atlantic and in the equatorial region of the western Atlantic, samples from the same depth as the maximum  $C^{14}$  layer yield  $C^{14}/C^{12}$  ratios identical with those of the North Atlantic Deep Water. It is possible that these areas lie beyond the influence of the boundary current.

Fifteen samples from the core of the NADW in both basins average  $-105$  per mil. No trend with latitude is apparent. Whereas the values for the lower Deep Water in the western trough agree with those for the overlying core water, the lower Deep Water in the eastern trough has a 27 per mil lower average than the average for

the overlying water. Although this difference suggests a longer residence time for water in the eastern basin, it is possible that it represents at least in part, contamination by water of Antarctic origin.

The South Atlantic samples corresponding to the Central Water mass (57, 58, 60 to 63) have  $C^{14}/C^{12}$  ratios about 10 per mil higher than the corresponding NACW and 5 per mil lower than the overlying surface water.

The Antarctic Intermediate Water in the South Atlantic is characterized by a relatively low  $C^{14}/C^{12}$  ratio. Six samples between the equator and  $40^\circ\text{S}$  average  $-118$  per mil.

Samples from  $20^\circ$  to  $40^\circ\text{S}$  (97 to 99) at depths between 1500 and 2200 meters show a significantly lower  $C^{14}/C^{12}$  ratio than is found farther north at similar depths. Because of the significantly lower salinities, this appears to be the result of mixing with lower-activity Antarctic water rather than of a longer residence time.

Antarctic Bottom Water of nearly pure type has been sampled in three places (135 to 137) in the southwestern Atlantic. Its average  $\Delta C^{14}$  value of  $-144$  per mil is the lowest found thus far in the Atlantic Ocean.

The radiocarbon results on a total of 78 surface samples from the Atlantic suggest that the samples taken within the core of a single water mass (as defined in oceanographic terms) show a standard deviation in  $\Delta C^{14}$  value only slightly exceeding the measurement error. Since significant differences exist between the various masses, it would appear that circulation within a water mass is relatively rapid with respect to the residence time of water particles in the mass.

*Discussion.* In order to obtain absolute mixing rate estimates from the radiocarbon data, a model of the operation of the oceans must be set up on the basis of existing oceanographic data. The resulting rates will be only as good as the model. In order to demonstrate the importance of the type of model adopted, two quite different sets of assumptions will be examined. In the first case the oceans are assumed to be at steady state (that is, the temperature, salinity,  $C^{14}/C^{12}$  ratio, etc., for any point in the ocean remains constant with time). The second model involves periods of rapid oceanic mixing separated in time by periods during which mixing

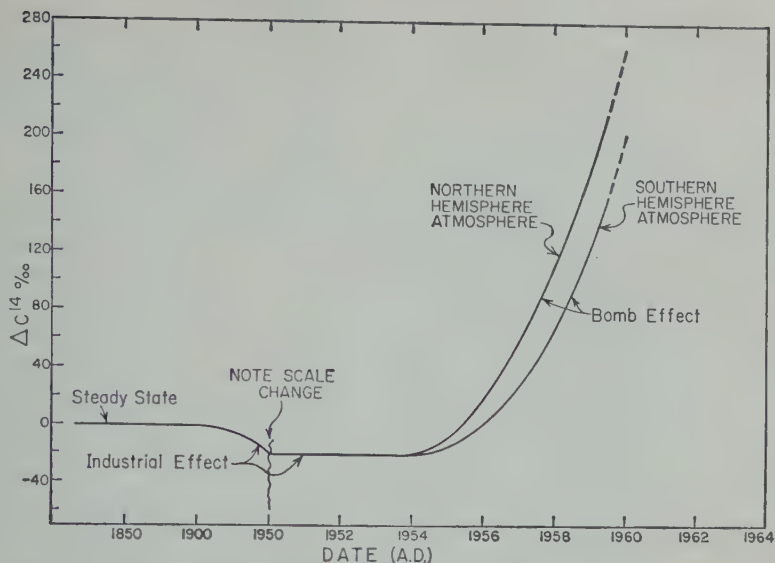


Fig. 5. Variation in the  $\Delta C^{14}$  value for atmospheric  $CO_2$  over the last century. The decrease beginning about 1900 is the result of the addition of industrial  $CO_2$ , and the rise beginning in 1954 is the result of  $C^{14}$  produced during nuclear tests.

is greatly restricted. In this case the properties characterizing any unit of volume in the system will change with time.

**Steady-state models.** Numerous investigators have shown that the  $C^{14}$  concentration in the atmosphere and surface water is presently not in steady state as the result of industrial  $CO_2$  and bomb-produced  $C^{14}$ . If calculations based on steady-state assumptions are to be made, it is necessary to determine what the concentrations of  $C^{14}$  in the samples measured would have been if the system had not been contaminated. The correction for bomb-produced  $C^{14}$  is accomplished as follows. Measurements of the  $C^{14}$  concentration in atmospheric  $CO_2$  and in vegetation define the rise in the  $C^{14}/C^{12}$  ratio for the atmospheric reservoir (see Fig. 5). Using the steady-state model described below it is possible to make successive approximations to estimate how much bomb-produced  $C^{14}$  was present in surface water at any given time. Such calculations indicate that samples collected before 1956 would contain negligible bomb-produced  $C^{14}$ . The results for such samples are given in Table 3, column 3, averaged on a geographical basis. The correction based on the model calculation has been made for each sample collected after 1956. The corrected averages for these samples are

given in column 4. Estimates of the prebomb concentrations based on results from samples taken both before and after January 1, 1956, are given in column 5. The  $C^{14}$  results of *Rafter and Fergusson* [1958] for the western South Pacific surface water and of *Fonselius and Ostlund* [1959] for the northern North Atlantic surface water have been treated in the same manner and are included for reference. These results should be strictly comparable, since interlaboratory calibrations have been carried out in each case (Table 3). Since the differences between the corrected (column 5) and uncorrected (column 2) averages are in all cases less than 1 per cent, the uncertainty in the corrections introduces no significant error into the calculations which follow. No correction has been made for samples from the deep water masses, as the model indicates that the amount of bomb-produced  $C^{14}$  added to these reservoirs would be negligible compared with that already present. For a more complete discussion of the quantity and distribution of bomb-produced  $C^{14}$  see *Broecker and Walton* [1959a] and *Broecker and Olson* [in press].

*Suess* [1955], *Hayes, Anderson, and Arnold* [1955], *Brannon, Daughtry, Perry, Whitaker, and Williams* [1957], *Fergusson* [1958], and



TABLE 2. Correction for Bomb-Produced Radiocarbon†† in Surface Ocean Water

| Geographic Region        | Uncorrected Avg. |     |       |        |  | Avg. pre-1/1/56 Samples§ |     |       |        |  | Avg. post-1/1/56 Samples |     |       |        |  | Corrected Avg. |     |       |        |  |
|--------------------------|------------------|-----|-------|--------|--|--------------------------|-----|-------|--------|--|--------------------------|-----|-------|--------|--|----------------|-----|-------|--------|--|
|                          | Avg.             | No. | S.D.¶ | S.E.** |  | Avg.                     | No. | S.D.¶ | S.E.** |  | Avg.                     | No. | S.D.¶ | S.E.** |  | Avg.           | No. | S.D.¶ | S.E.** |  |
| North Atlantic† 60°-80°N | -35              | 3   | ±2    | ±4     |  | ...                      | 0   | ...   | ...    |  | -38                      | 3   | ±2    | ±4     |  | -38            | 3   | ±2    | ±4     |  |
| North Atlantic* 15°-40°N | -49              | 18  | ±10   | ±3     |  | -49                      | 9   | ±8    | ±3     |  | -54                      | 9   | ±9    | ±3     |  | -52            | 18  | ±9    | ±2     |  |
| Caribbean*               | -56              | 8   | ±9    | ±3     |  | -56                      | 8   | ±9    | ±3     |  | ...                      | 0   | ...   | ...    |  | -56            | 8   | ±9    | ±3     |  |
| South Atlantic* 0°-40°S  | -57              | 16  | ±10   | ±3     |  | ...                      | 0   | ...   | ...    |  | -63                      | 16  | ±10   | ±3     |  | -63            | 16  | ±10   | ±3     |  |
| Falkland Current         | -77              | 3   | ±3    | ±4     |  | ...                      | 0   | ...   | ...    |  | -80                      | 3   | ±3    | ±4     |  | -80            | 3   | ±3    | ±4     |  |
| Antarctic†*              |                  |     |       |        |  |                          |     |       |        |  |                          |     |       |        |  |                |     |       |        |  |
| (temp. <3°C)             | -120             | 3   | ±10   | ±6     |  | ...                      | 0   | ...   | ...    |  | -123                     | 3   | ±10   | ±6     |  | -123           | 3   | ±10   | ±6     |  |
| South Pacific† 15°-42°S  | -41              | 16  | ±8    | ±2     |  | -54                      | 2   | ±2    | ±3     |  | -48                      | 14  | ±6    | ±2     |  | -49            | 16  | ±6    | ±2     |  |

\* Based on results given in this paper.  
† Based on results published by *Rafter and Fergusson* [1958].  
‡ Based on results published by *Fonselius and Ostlund* [1959].  
§ Uncorrected (bomb-produced C14 assumed negligible).  
|| Corrected for bomb-produced C14 using steady-state circulation model.  
¶ Standard deviation of results from mean.  
\*\* Standard error on the mean (based on the larger of the S.D. or average laboratory error).  
†† All results given in ΔC14 units (‰).

LE 3. Interlaboratory Calibration (The results are given on the  $\Delta C^{14}$  scale as per mil difference from 0.950 the activity of the N.B.S. oxalic acid standard used as a primary standard by all three laboratories.)\*

|                         | Lamont             | New Zealand | Stockholm |
|-------------------------|--------------------|-------------|-----------|
| Untreated wood          |                    |             |           |
| Standard                |                    |             |           |
| 1890 oak                | $-6 \pm 5^\dagger$ | ...         | ...       |
| New Zealand wood        |                    |             |           |
| Standard (contemporary) |                    |             |           |
| 1890 oak                | $-18 \pm 6$        | $-20 \pm 4$ | ...       |
| Stockholm wood          | $8 \pm 6$          | ...         | $6 \pm 5$ |
| Standard (1850 oak)     |                    |             |           |

The following relationships have been used to convert the data on ocean water samples given by *Strogonov and Fergusson* [1958] and *Fonselius and Stroecker* [1959] to the  $\Delta C^{14}$  scale.

$$R = 10R + 6.0 \quad (\text{Stockholm results})$$

$$= 10 \left[ \frac{\% C^{14} \text{ enrichment}}{\text{W.R.T. N.Z. wood standard}} \right] - 66.0 \quad (\text{New Zealand results})$$

This result supersedes that of  $0 \pm 5$  published by *Stroecker and Olson* [1959]. The change is the result of further comparison measurements.

*Stroecker and Vogel* [1958] have published data indicating the decrease in the atmospheric  $C^{14}/C^{12}$  ratio as the result of the consumption of fossils. The combined results suggest a  $25 \pm 5$  per mil reduction in the average atmospheric  $C^{14}/C^{12}$  ratio between 1855 and 1955. Measurements of comparable accuracy are not available for the surface oceans. Again the steady-state model discussed below has been used to estimate the magnitude of the oceanic reduction. Such calculations suggest that the reduction in the  $C^{14}/C^{12}$  ratio in the surface 100 meters of the Atlantic and Pacific oceans should be  $0.67$  of that for the atmosphere. A value of  $0.40 \pm .20$  for the atmospheric reduction is estimated for vertically mixed regions at the northern and southern extremes of the ocean. The preindustrial  $\Delta C^{14}$  values estimated for various areas on this basis are listed in Table 4. Assuming that the entire surface of the Pacific and Indian Oceans has  $\Delta C^{14}$  values averaging similar to those of *Rafter and Fergusson* [1958]

for the southwestern Pacific, a world-average  $\Delta C^{14}$  value has been computed for surface ocean water in 1955 and in 1855. Again, for the deep water masses the reduction in the  $C^{14}/C^{12}$  ratio as a result of equilibration with 'fossil'  $CO_2$  can be considered negligible in most cases.

Estimates of the mean residence times of water molecules in the subsurface water masses of the Atlantic Ocean can easily be made if the following assumptions are valid: (1) circulation in the oceans is at steady state; (2) the corrected  $\Delta C^{14}$  value for surface water in the convergence (the area considered on the basis of oceanographic evidence to be the source region for a given subsurface water mass) characterizes the 'new' water continually being added to the mass; (3) mixing between adjacent subsurface water masses does not measurably alter their respective  $C^{14}/C^{12}$  ratios; (4)  $CO_2$  supplied to a subsurface mass through solution of carbonate sediment and through the oxidation of organic material does not measurably alter the  $C^{14}/C^{12}$  ratio of the  $CO_2$  initially present in the mass. In this case, in the age range of 0 to 1000 years,

$$\tau_M = 8.04 \left[ (\Delta C^{14}_{\text{surface}} - \Delta C^{14}_{\text{depth}}) - \frac{(\Delta C^{14*}_{\text{surface}} - \Delta C^{14*}_{\text{depth}})}{2000} \right] \text{ yr}$$

where  $\tau_M$  is the mean residence time of  $H_2O$  molecules within a given water mass (in other words, the time required for a volume of water equal to that contained in the deep water mass to be supplied from the surface source region).

The results of computations made on this basis for the various subsurface water masses of the Atlantic Ocean are given in Table 5. The locations of the source areas for the individual water masses are those given in *Sverdrup, Johnson, and Fleming* [1942]. The errors quoted are based on the uncertainty in the measurement averages only and hence do not include the uncertainty in the assumptions upon which the calculations are based.

The results suggest the following conclusions:

1. Water masses originating in the Antarctic are in general characterized by shorter residence times than those originating in the high latitudes of the North Atlantic.

TABLE 4. Correction for Industrial CO<sub>2</sub> Effect in Surface Oceans

| Geographic Location        | Average $\Delta C^{14}$<br>Surface Oceans<br>1955<br>(see Table 2),<br>‰ | Fraction Atmospheric<br>Industrial CO <sub>2</sub> Effect<br>(based on model<br>calculations) | Estimated $\Delta C^{14}$<br>Reduction from<br>Industrial CO <sub>2</sub> *<br>(1855 to 1955),<br>‰ | Average $\Delta C^{14}$<br>Surface Oceans<br>1855<br>(calculated),<br>‰ |
|----------------------------|--|---|---|---|
| North Atlantic<br>60°-80°N | -38 ± 4  | .40 ± .20   | 10 ± 8  | -28 ± 9   |
| North Atlantic<br>15°-40°N | -52 ± 2  | .67 ± .15   | 17 ± 7  | -35 ± 7   |
| Caribbean Sea              | -56 ± 3  | .67 ± .15   | 17 ± 7  | -39 ± 8   |
| South Atlantic<br>0°-40°S  | -63 ± 3  | .67 ± .15   | 17 ± 7  | -46 ± 8   |
| Falkland Current           | -80 ± 4  | .40 ± .15   | 10 ± 8  | -70 ± 9   |
| Antarctic <3°C             | -123 ± 6   | .40 ± .25   | 10 ± 8  | -113 ± 10   |
| South Pacific<br>15°-42°S  | -49 ± 2  | .67 ± .15   | 17 ± 7  | -32 ± 7   |
| Average surface ocean      | -58 ± 10   | .60 ± .15   | 15 ± 8  | -43 ± 13  |

\* Assumes industrial CO<sub>2</sub> effect in atmosphere of -25 ± 10 per mil in 1955.

2. Except for the NACW and for the wedge of water lying at depths between 1200 and 2400 meters in the western North Atlantic, most of the subsurface water of North Atlantic origin appears to be at least 600 years 'old.'

3. Whereas the deepest waters in the western basin of the North Atlantic have the same residence time as the overlying water, the water below 4000 meters in the eastern basin appears to have a significantly higher residence time (900 years as opposed to ~ 650 years).

4. Samples from the SACW appear to have  $\Delta C^{14}$  values nearly identical to those for surface water samples taken in the outcrop area of the water mass. This suggests that the water is in fairly rapid communication with the surface. The decrease in  $\Delta C^{14}$  toward the bottom of the mass may be the result of mixing with the low- $C^{14}$  water of the AAIW mass either at depth or in the outcrop region. In this case, the added uncertainty in the relative correction for the industrial effect for the surface and subsurface masses becomes important, yielding a possible range of apparent ages from 0 to 200 years.

Of the four assumptions on which these residence times are based only the fourth can be adequately verified. The amount of CO<sub>2</sub> added to the subsurface water masses of the Atlantic

through oxidation of organic material does not exceed 10 per cent of the CO<sub>2</sub> initially present. If the oxidized organic material were carried along with the surface water entering the mass it would probably have the same  $\Delta C^{14}$  value as the water and hence would have no effect. On the other hand, the organic material settles downward rapidly from the surface water, the CO<sub>2</sub> produced by oxidation would have a  $\Delta C^{14}$  value significantly higher than that of the deep water to which it was added—as much as 80 per mil in the North Atlantic and 80 per mil in the South Atlantic, assuming that oxygen is being consumed at a linear rate. The maximum for the South Atlantic exceeds that for the North Atlantic because of the shorter residence times and the greater surface-depth contrast in  $\Delta C^{14}$  in the South Atlantic. Thus for the worst case (that is, if all the organic material oxidized in a deep water mass has the  $\Delta C^{14}$  of the overlying surface water), the residence times quoted would have to be increased by 30 years in the North Atlantic and 70 years in the South Atlantic.

The steady-state rate at which carbonates in sediments would have to be dissolved and/or exchanged in order measurably to alter the C<sup>13</sup>/C<sup>12</sup> ratio in near-bottom water masses is prohibitively large, exceeding the average rate

TABLE 5. Mean Residence Time Estimates for Subsurface Water Masses of the Atlantic Ocean (subject to validity of assumptions stated in text)

| Location and Depth   | C <sup>14</sup> Results on<br>Samples from within Mass |     |            |            | Index No. of<br>Samples Used in<br>Calculation of<br>Averages | Estimated<br>Pre-Industrial<br>Revolution<br>$\Delta C^{14}$ Value for<br>Source Area,<br>‰ | M,<br>years |
|--|--|-----|------------|------------|---|---|-------------|
|  | Average<br>$\Delta C^{14}$ , ‰                         | No. | S.D.,<br>‰ | S.E.,<br>‰ |   |   |             |
| North Atlantic<br>30°N, 200-400 m                                    | -71  | 3   | ±4         | ±4         | 52-54   | -35 ± 7   | 300 ± 70    |
| North Atlantic<br>25°N, 800-1100 m                                   | -99  | 8   | ±11        | ±4         | 65-72   | -28 ± 9   | 640 ± 90    |
| Central North Atlantic<br>40°N,<br>00-2400 m                         | -72  | 8   | ±16        | ±6         | 80-84, 87-89  | -28 ± 9   | 380 ± 100   |
| Central Atlantic<br>40°W, 20°S-25°N,<br>00-2000 m                    | -104   | 7   | ±8         | ±3         | 90-96   | -28 ± 9   | 690 ± 90    |
| South Atlantic Deep<br>40°S-35°N,<br>00-4000 m                       | -105   | 15  | ±11        | ±3         | 100-114   | -28 ± 9   | 700 ± 90*   |
| South Atlantic Bottom<br>Water 23°-40°N,<br>Western Basin,<br>4000 m | -104 (-100)†   | 16  | ±13        | ±3         | 115-127, 130, 131   | -28 ± 9   | 690 ± 90    |
| South Atlantic Bottom<br>Water 15°-25°N,<br>Western Basin,<br>4000 m | -127 (-123)†   | 3   | ±8         | ±5         | 129, 133, 134   | -28 ± 9   | 920 ± 100   |
| South Atlantic<br>0°S, 200-400 m                                     | -61  | 6   | ±12        | ±5         | 57, 58, 60-63   | -46 ± 8   | <250        |
| South Atlantic<br>0°S, 500-1200 m                                    | -118   | 6   | ±17        | ±7         | 73-78   | -90 ± 20  | <350*       |
| South Atlantic<br>0°S-20°S,<br>00-2200 m                             | -130   | 3   | ±12        | ±7         | 97-99   | ...   | ...         |
| Arctic Bottom Water<br>0°S, >4000 m                                  | -144   | 3   | ±10        | ±6         | 135-137   | -113 ± 15   | <350*       |

These averages represent the cores of well-defined water masses; NADW, AAIW, and AABW. Corrected for the contribution of AABW (based on estimates given by *Wüst*, 1957b).

accumulation of carbonate oozes. Deep sea cores show no evidence of large-scale carbonate solution and the possibility of sufficiently rapid exchange between the carbonates in the sediments and the overlying water is excluded by the fact that the C<sup>14</sup>-age and O<sup>18</sup>-temperature records are preserved in deep sea carbonates. From these numerous other lines of evidence it can be inferred that the presence of carbonate sediments does not measurably influence the residence time computed for the overlying water mass. In red sea areas carbonate falling from the surface is presumably redissolved before or soon after

reaching the bottom. In this case the CO<sub>2</sub> added to the water through carbonate solution may well be higher in C<sup>14</sup> than the water itself. Using an addition rate comparable to the sedimentation rate of carbonate in globigerina ooze areas (~ 1 g/cm<sup>2</sup>/1000 yr) and assuming that the dissolved carbonate is mixed throughout the lower 100 meters of ocean water, the addition per 100 years would be equivalent to 5.0 per cent of the total CO<sub>2</sub> already present in the water. If this is the case the contribution to the apparent age of the water from this source is negligible.



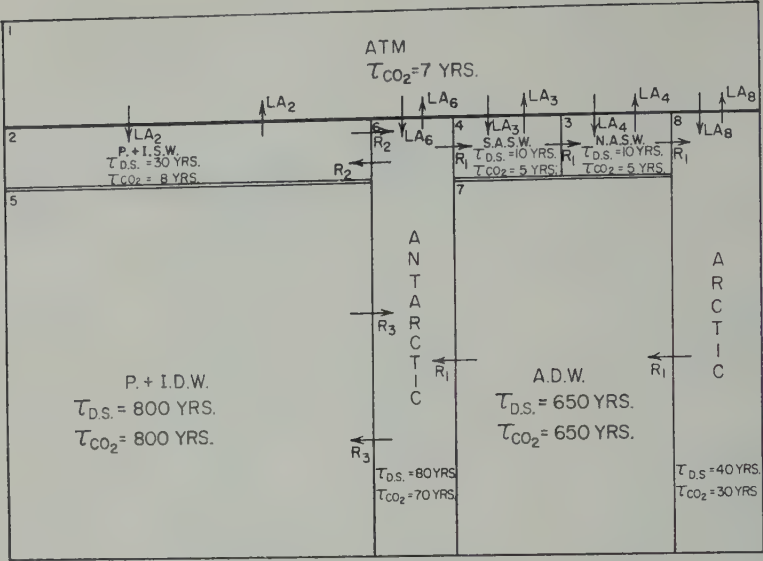


Fig. 6. Model of steady-state CO<sub>2</sub> cycle in the atmosphere and oceans. The numbers in the upper left-hand corner of each box identify the reservoir. The arrows indicate possible modes of transfer between reservoirs, and the letters show the rates of transfer. The CO<sub>2</sub> and DS (dissolved solid) mean residence times differ because significant quantities of CO<sub>2</sub> are transferred through the atmosphere.

In the case of the samples from the cores of the NADW, AAIW, and AABW, complications resulting from mixing at depth are at a minimum. For the other entries in Table 5 the possible effects of mixing must be considered. For example, in the case of the central water systems, where vertical temperature gradients suggest mixing of water of more than one source, it is possible that ages computed in the above manner are in error. Thus, as suggested by *Bowen and Sugihara* [1960], the 300 year age computed for the NACW may be too great. The low  $\Delta C^{14}$  value could, as they propose, be the result of mixing with the underlying low  $\Delta C^{14}$  NADW mass. Certainly the natural radiocarbon method is more valuable in the study of deep water masses than in the study of near-surface masses where mixing is serious and where natural radioactive tracers of shorter half-life or man-made radioactive tracers may be used.

Mixing between the NADW and the AABW poses a problem when bottom water samples from the Atlantic are analyzed. From potential temperature and salinity data *Wüst* [1957b] estimated the relative proportions of each type of water at various locations in the Atlantic. Using *Wüst's* estimates the results for bottom

samples primarily of North Atlantic origin have been corrected for the decrease in  $C^{14}/C^{12}$  ratio introduced by the admixture of AABW. These corrections range from 2 to 5 per mil with one (index No. 128) as high as 8 per mil. As shown in Table 5, correction for the contribution of AABW changes the bottom water averages in both basins by only 4 per mil.

The validity of the second as well as the first assumption depends on how close the ocean is to steady-state circulation. Discussion of this point will be deferred to a later section of the paper.

If the ocean is at steady state it should be possible to construct a circulation model consistent with oceanographic evidence which would yield the observed distribution of  $C^{14}$  in the ocean-atmosphere system. Such models have been discussed by *Arnold and Anderson* [1957], *Suess and Revelle* [1957], *Craig* [1957a], *Brammer, Daughtry, Perry, Whitaker, and Williams* [1957]. The present model is designed to incorporate more details of the system than could be handled by the three-layer (atmosphere, surface ocean, deep ocean) models used by the above authors without greatly increasing the number of unknown parameters. As is shown in

TABLE 6. Parameters Defining Reservoirs for Steady-State Ocean-Atmosphere Model

| Reservoir No. | Reservoir Name                 | Latitude Range | Depth Range, meters | Surface Area, meter <sup>2</sup> | CO <sub>2</sub> Content, moles | $\Delta C^{14}$ , ‰ |
|---------------|--------------------------------|----------------|---------------------|----------------------------------|--------------------------------|---------------------|
| 1             | Atmosphere                     | Entire         | Entire              | $360 \times 10^{12}$             | $57 \times 10^{15}$            | 0                   |
| 2             | Pacific + Indian Surface Water | <50°S          | <100                | $220 \times 10^{12}$             | $50 \times 10^{15}$            | -32                 |
| 3             | South Atlantic Surface Water   | 50°S-5°N       | <100                | $40 \times 10^{12}$              | $9 \times 10^{15}$             | ...                 |
| 4             | North Atlantic Surface Water   | 5°N-55°N       | <100                | $40 \times 10^{12}$              | $9 \times 10^{15}$             | -35                 |
| 5             | Pacific + Indian Deep Water    | <50°S          | >100                | ...                              | $1980 \times 10^{15}$          | -200*               |
| 6             | Antarctic                      | >50°S          | Entire              | $45 \times 10^{12}$              | $420 \times 10^{15}$           | ...                 |
| 7             | Atlantic Deep Water            | 50°S-55°N      | >100                | ...                              | $655 \times 10^{15}$           | -104                |
| 8             | Arctic                         | >55°N          | Entire              | $15 \times 10^{12}$              | $40 \times 10^{15}$            | ...                 |
| T             | Reservoirs 2-8                 | Entire         | Entire              | $360 \times 10^{12}$             | $3220 \times 10^{15}$          | -160                |

Based on results given by *Rafter and Fergusson* [1958], and *Suess, Rakestraw, and Oeschger* [1959] and two results obtained at Lamont on deep water samples collected by the University of Washington off the Aleutians ( $\Delta C^{14} = -233 \pm 7$  and  $-210 \pm 9$ ).

Figure 6, the ocean has been divided into seven internally homogeneous reservoirs. Mixing between the reservoirs proceeds only as is indicated by the arrows. The Atlantic Ocean is characterized by a northward transfer of surface water from the Antarctic to the South Atlantic to the North Atlantic to the Arctic, balanced by a return flow from the Arctic to the Atlantic through the NADW system. The Atlantic deep water masses of Antarctic origin are considered to be part of the Antarctic reservoir. Because only limited data are available for the Pacific and Indian oceans a simple two-fold division has been made—one surface and one deep water reservoir. Mixing across the main thermocline in the North and South Atlantic and Pacific-Indian surface reservoirs is assumed to be negligible, the deep water masses being fed from locally mixed reservoirs at the extremes of the ocean. Transfer from the Pacific to the Atlantic reservoir through the Bering Strait is assumed to be negligible. Transfer of CO<sub>2</sub> from the ocean to the atmosphere is assumed to occur at a uniform rate everywhere. The locations and depth ranges are tabulated for each reservoir (Table 6).

At steady state the  $\Delta C^{14}$  values for the various reservoirs are governed by the following equations:

$$LA_1 \left( \frac{\Delta C^{14}_1 - \Delta C^{14}_{ASW}}{1000} \right) = \lambda (T_2 + T_3 + T_4 + T_5 + T_6 + T_7 + T_8) \times \left( 1 + \frac{\Delta C^{14}_T}{1000} \right) \quad \text{Reservoir 1}$$

$$LA_2 \left( \frac{\Delta C^{14}_1 - \Delta C^{14}_2}{1000} \right) + R_2 \left( \frac{\Delta C^{14}_6 - \Delta C^{14}_2}{1000} \right) = T_2 \lambda \left( 1 + \frac{\Delta C^{14}_2}{1000} \right) \quad \text{Reservoir 2}$$

$$LA_3 \left( \frac{\Delta C^{14}_1 - \Delta C^{14}_3}{1000} \right) + R_1 \left( \frac{\Delta C^{14}_6 - \Delta C^{14}_3}{1000} \right) = T_3 \lambda \left( 1 + \frac{\Delta C^{14}_3}{1000} \right) \quad \text{Reservoir 3}$$

$$LA_4 \left( \frac{\Delta C^{14}_1 - \Delta C^{14}_4}{1000} \right) + R_1 \left( \frac{\Delta C^{14}_3 - \Delta C^{14}_4}{1000} \right) = T_4 \lambda \left( 1 + \frac{\Delta C^{14}_4}{1000} \right) \quad \text{Reservoir 4}$$

$$R_3 \left( \frac{\Delta C^{14}_6 - \Delta C^{14}_5}{1000} \right) = T_5 \lambda \left( 1 + \frac{\Delta C^{14}_5}{1000} \right) \quad \text{Reservoir 5}$$

$$\begin{aligned} &L A_6 \left( \frac{\Delta C^{14}_1 - \Delta C^{14}_6}{1000} \right) + R_2 \left( \frac{\Delta C^{14}_2 - \Delta C^{14}_6}{1000} \right) \\ &\quad + R_3 \left( \frac{\Delta C^{14}_5 - \Delta C^{14}_6}{1000} \right) \\ &\quad + R_1 \left( \frac{\Delta C^{14}_7 - \Delta C^{14}_6}{1000} \right) \\ &= T_6 \lambda \left( 1 + \frac{\Delta C^{14}_6}{1000} \right) \quad \text{Reservoir 6} \\ &R_1 \left( \frac{\Delta C^{14}_8 - \Delta C^{14}_7}{1000} \right) \\ &= T_7 \lambda \left( 1 + \frac{\Delta C^{14}_7}{1000} \right) \quad \text{Reservoir 7} \\ &L A_8 \left( \frac{\Delta C^{14}_1 - \Delta C^{14}_8}{1000} \right) + R_1 \left( \frac{\Delta C^{14}_4 - \Delta C^{14}_8}{1000} \right) \\ &= T_8 \lambda \left( 1 + \frac{\Delta C^{14}_8}{1000} \right) \quad \text{Reservoir 8} \end{aligned}$$

where  $L$  is the rate of exchange of  $\text{CO}_2$  between the atmosphere and surface ocean (in moles  $\text{CO}_2/\text{meter}^2/\text{year}$ ),  $A_i$  is the area a given reservoir exposes to the atmosphere (in  $\text{meter}^2$ ),  $T_i$  is

the quantity of  $\text{CO}_2$  contained in a given reservoir (in moles  $\text{CO}_2$ ),  $R_i$  is the rate of transfers of  $\text{CO}_2$  from one reservoir to another (in moles  $\text{CO}_2/\text{year}$ ), and  $\lambda$  is the decay constant of  $\text{C}^{14}$  (in  $\text{years}^{-1}$ ).

Of these eight equations seven are linearly independent, allowing seven unknowns to be determined ( $L$ ,  $R_1$ ,  $R_2$ ,  $R_3$ ,  $\Delta C^{14}_4$ ,  $\Delta C^{14}_6$ , and  $\Delta C^{14}_8$ ). The values selected for the other parameters appearing in the equation are given in Table 6. The values computed for the unknowns are given in Table 7. The value of  $L$  is independent of the arrangement of the reservoirs or the mode of mixing between them. It depends only on the average  $\text{C}^{14}/\text{C}^{12}$  ratio of ocean water,  $\Delta C^{14}_T$ , and the difference between the average  $\text{C}^{14}$  concentration in atmospheric and surface ocean  $\text{CO}_2$ . Since the difference between  $\Delta C^{14}_T$  and  $\Delta C^{14}_{\text{ASW}}$  is known to an accuracy of perhaps 25 per cent and since  $\Delta C^{14}_T$  is known to an accuracy of about 10 per cent, the estimate of  $L$  should be uncertain by no more than 30 per cent. The result confirms similar estimates of this parameter obtained by Craig [1957a] and is also in satisfactory agreement with the value of 5 moles/ $\text{m}^2/\text{yr}$  obtained by Broecker and Walton [1959b] for four closed basin lakes in the

TABLE 7. Comparison of Results of Steady-State Model Calculations with Independent Estimates

| Parameter         | Description   | Value Based on Model                                       | Independent Estimate  | Method                      |
|-------------------|---|--|---|-----------------------------|
| $\Delta C^{14}_4$ | $\text{C}^{14}$ concentration in South Atlantic Surface Water | -65 %  | -46 %*  | Direct measurement          |
| $\Delta C^{14}_6$ | $\text{C}^{14}$ concentration in the Antarctic reservoir      | -120 %   | -113 %*   | Direct measurement          |
| $\Delta C^{14}_8$ | $\text{C}^{14}$ concentration in the Arctic reservoir         | -30 %  | -28 %*  | Direct measurement          |
| $L$               | Ocean-atmosphere exchange rate                                | $22 \pm 7 \frac{\text{moles CO}_2}{\text{m}^2 \text{ yr}}$ | $5 \frac{\text{moles CO}_2^\dagger}{\text{m}^2 \text{ yr}}$       | Inland lakes                |
| $R_1$             | Rate of transfer between reservoirs in Atlantic cycle         | $1.0 \times 10^{15} \frac{\text{moles CO}_2}{\text{yr}}$   | $0.4 \times 10^{15} \frac{\text{moles CO}_2^\ddagger}{\text{yr}}$ | Surface current measurement |
| $R_2$             | Rate of exchange between P. + I. Surface Water and Antarctic  | $1.6 \times 10^{15} \frac{\text{moles CO}_2}{\text{yr}}$   | ...   | ...                         |
| $R_3$             | Rate of exchange between P. + I. Deep Water and Antarctic     | $2.5 \times 10^{15} \frac{\text{moles CO}_2}{\text{yr}}$   | ...   | ...                         |

\* See Table 4.

† Broecker and Walton [1959b] and Broecker and Olson [in press].

‡ Based on net northward rate of surface water transport given by Sverdrup, Johnson, and Fleming [1942]

t Basin region of the United States. The estimates are based on the steady-state concentration in the lake-river systems. Independent confirmation has been obtained in case (Great Salt Lake) on the basis of the uptake of bomb-produced  $C^{14}$  [Broecker Olson, in press]. The higher value for the ocean probably reflects a higher degree of surface agitation and hence mixing.

The  $\Delta^{14}$  values predicted by the model for Antarctic and Arctic are, as shown in Table 1, in good agreement with estimates based on measurements in these areas (see *Fonselius and Ostlund* [1959], *Rafter and Ferguson* [1958], and this paper). The disagreement between the predicted and measured values for the South Atlantic could well indicate back-mixing from the North to the South Atlantic surface system. Agreement could be achieved by imposing an exchange rate between these reservoirs of about  $1.0 \times 10^{15}$  moles  $CO_2$ /yr on the net northward transfer rate of  $1.0 \times 10^{15}$  moles  $CO_2$ /yr. Since the one-way mixing cycle time for the Atlantic is certainly an overestimation, a more adequate model would have to take into account back-mixing between pairs of reservoirs. Too many additional parameters are required for any realistic estimates of the magnitude of these effects to be made.

It is of interest to compare net rates of northward transport based on oceanographic data with that of  $1.0 \times 10^{15}$  moles  $CO_2$  per year predicted by the steady-state model. *Sverdrup, Johnson, and Fleming* [1942] gave an estimate equivalent to  $0.4 \times 10^{15}$  moles  $CO_2$  per year for surface currents, and *Wüst* [1957a, 1957b] gave a value equivalent to  $0.8 \times 10^{15}$  moles of  $CO_2$  per year for the upper 200 meters in the South Atlantic.

From these results the mean residence times for both  $CO_2$  and dissolved solids (DS) in the various reservoirs can be computed. The residence times given in Figure 6 are computed by dividing the total amount of  $CO_2$  (or DS) in the reservoir by the rate at which  $CO_2$  (or DS) is being added to the reservoir. It is assumed that atmospheric transfer of dissolved solids is negligible compared with oceanic transport.

Assuming steady state, *Stommel* [1957b] calculated, on the basis of estimates of the quantity of water transported by the Western Boundary Current and by the Antarctic Bottom Water, that the mean residence time for water below 2000 meters should be about 500 years. This estimate agrees favorably with that based on the  $C^{14}$  data.

It is of interest to note that, whereas the  $C^{14}/C^{13}$  ratio for much of the deep Pacific is about 100 per mil lower than that for the NADW, the

TABLE 8. Comparison of Oceanic Tritium Concentrations Predicted on the Basis of the Steady-State Model with Measured Values\*

| Reservoir No.<br>(see Fig. 6) | Steady-State Tritium<br>Conc. (predicted if<br>production rate =<br>$0.5T/cm^2/sec$ ), T.U. | Steady-State Tritium<br>Conc. (predicted if<br>production rate =<br>$1.5T/cm^2/sec$ ), T.U. | Steady-State Tritium<br>Concentration<br>(measured), T.U. |
|-------------------------------|---|---|---|
| 2                             | 0.4   | 1.1   | 1.2 (2) S.D. = .4   |
| 3                             | 0.2   | 0.6   | ...   |
| 4                             | 0.3   | 1.0   | 1.2 (11) S.D. = .4  |
| 5                             | $0.7 \times 10^{-3}$  | $2 \times 10^{-3}$  | ...   |
| 6                             | 0.03  | 0.1   | ...   |
| 7                             | $3 \times 10^{-3}$  | $10 \times 10^{-3}$   | <0.1 (3)  |
| 8                             | 0.1   | 0.4   | ...   |

\* Assumes tritium production rate such that between 0.5 and 1.5 tritium atoms/ $cm^2/sec$  reach the earth's surface (this range includes the estimates made by *Begamann and Libby* [1957] ( $1.0 T \text{ atoms}/cm^2/sec$ ), *Giletti* [1957a] ( $1.2 \pm .5 T \text{ atoms}/cm^2/sec$ ), *Giletti, Bazan, and Kulp* [1958] ( $.75 \pm .4 T \text{ atoms}/cm^2/sec$ ), *Giletti* [1958] ( $.9 \pm .2 T \text{ atoms}/cm^2/sec$ ).

\* Based on data published by *Giletti* [1958] and by *Begamann and Libby* [1957] (all samples included collected before March 1954).



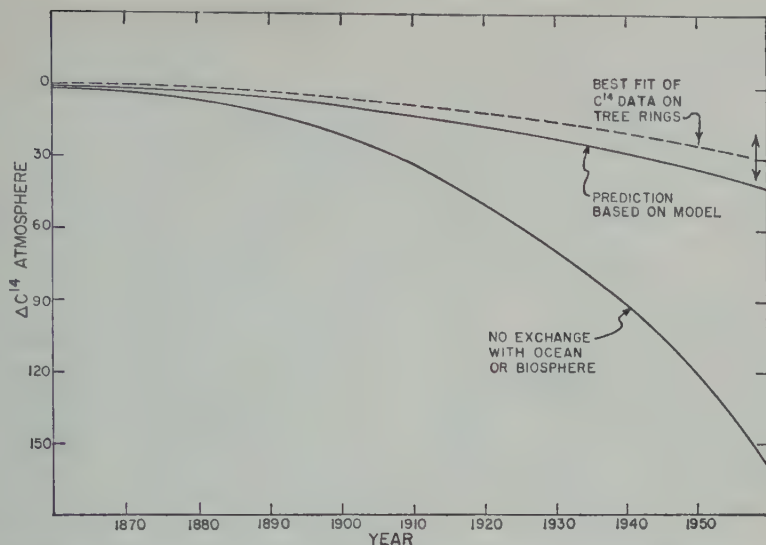
computed residence time is only slightly larger. This is true because the source for the deep Pacific has approximately a 90 per mil lower  $C^{14}/C^{12}$  ratio than that for the NADW.

One potential test (first suggested by Craig [1957b]) is to determine whether the distribution of natural tritium in the oceans is consistent with the rates of mixing fixed by the  $C^{14}$  distribution. Because no wide coverage of the tritium concentration in the ocean is available for the period prior to the onset of large-scale bomb-production of tritium (March 1954), the comparison is made as follows. Using extreme values of 0.5 and 1.5 T atoms/cm<sup>2</sup>/sec for the rate at which tritium is added to the troposphere (based on the estimates given by Begamann and Libby [1957], 1.0 T atoms/cm<sup>2</sup>/sec; Craig [1957b], 1.2 T atoms/cm<sup>2</sup>/sec; Giletti, Bazan, and Kulp [1958], 0.75 T atoms/cm<sup>2</sup>/sec and Bolin [1958],  $0.9 \pm 0.2$  T atoms/cm<sup>2</sup>/sec), an upper and lower limit for the tritium concentration in the various oceanic reservoirs can be made. Three assumptions are required: (1) of the total tritium added to the troposphere, 90 per cent ultimately reaches and decays in the oceans; (2) tritium is added to the ocean in a geographically uniform manner; (3) the tritium cycle was at steady state prior to 1954.

The results of this calculation are given in Table 8 along with averages of the existing data. For both the North Atlantic and the Pacific-Indian surface reservoirs the measured values fall just beyond the allowed range. Because of the uncertainties in the production rate and in assumptions 2 and 3, the agreement may be considered satisfactory, suggesting that the thicknesses and residence times of water in the surface reservoirs are realistic.<sup>2</sup>

<sup>2</sup> Since the preparation of this paper, Craig and Lal (private communication) have suggested that the tritium production rate is actually as low as 0.35 T atoms/cm<sup>2</sup>/sec. If this is the case, it is difficult to explain the high concentrations of tritium found in the surface ocean by Begamann and Libby [1957] and by Giletti, Bazan, and Kulp [1958]. The lack of agreement between the production rate estimates and the surface inventories is made far worse if vertical mixing between the surface and central waters is as rapid as would be suggested by Bowen and Sugihara's [1960] Sr<sup>90</sup> data. The uncertainties introduced by these recent developments remove the value of tritium data as a check on the circulation model proposed here.

A second independent test of the model is to compare the measured reduction in the  $C^{13}/C^{12}$  ratio in atmospheric CO<sub>2</sub> resulting from the addition of fossil CO<sub>2</sub> to the system with the reduction predicted by the steady-state model. The predicted decrease is based on the estimate of fossil CO<sub>2</sub> production given by Suess and Revelle [1957] for 10-year periods beginning in 1860 and on the following assumptions: (1) the excess CO<sub>2</sub> added to the system does not measurably ( $\pm 10$  per cent) alter the transfer rate of CO<sub>2</sub> between the ocean and atmosphere; (2) the CO<sub>2</sub> in the atmosphere equilibrates isotopically with the living terrestrial biosphere within a few years but does not equilibrate measurably ( $\pm 10$  per cent) with the humus of terrestrial soils in the time available ( $\sim 3$  years). Because the seven simultaneous differential equations involved in this calculation are not readily solved, the calculation has been carried out numerically. The rate at which CO<sub>2</sub> is added instantaneously to one reservoir (the atmosphere) mixes into the other reservoirs in a time first determined. The fossil CO<sub>2</sub> may be considered a series of negative  $C^{14}$  spikes added to the system at various times between 1860 and 1955. The degree of mixing of each of these spikes can then be determined for any given time after addition. The sum of the contribution of each spike yields the net effect in a given reservoir for any given time between 1860 and 1955. The results are plotted in Figure 7 along with the best estimate of the actual reduction based on  $C^{14}$  measurements on tree rings [Suess, 1955; Brannon, Daughtry, Perry, Whitaker, and Williams, 1957; Fergusson, 1958; Munnich and Vogel, 1958; Hayes, Anderson, and Arnold, 1955]. Whereas Fergusson's data appear to be the most definitive, the consistently higher estimates obtained by the other authors have been given weight by increasing all Fergusson's reduction estimates by 25 per cent. The double arrow in the diagram represents the uncertainty in the estimates based on tree ring data. The curve for no exchange between the fossil CO<sub>2</sub> and the carbon of reservoirs other than the atmosphere is given for comparison. Whereas the predicted curve yields a somewhat greater reduction than has been observed, the two estimates are sufficiently close so that the difference could result from errors in the assumptions (particularly the one con-



g. 7. Changes in the  $C^{14}/C^{12}$  ratios of atmospheric  $CO_2$  resulting from dilution with industrial  $CO_2$ . The dashed curve is based on the  $25 \pm 10$  per mil reduction suggested by tree ring data (see text), and the upper solid curve is based on  $CO_2$  production estimates and the steady-state model. The double arrow indicates an estimate of the uncertainty in the experimental curve.

ing the exchange with the humus reservoir) rather than weaknesses in the steady-state model. In the thickness of and residence time of carbon in the surface reservoirs largely control the prediction.

Polin and Eriksson [1959] recently made a similar prediction of the effect of fossil  $CO_2$  on the  $C^{14}/C^{12}$  ratio in atmospheric  $CO_2$  by using a three-layer model referred to above. Selecting values of their parameters  $\tau_{A-M}$  and  $\tau_{D-M}$  corresponding to a  $\Delta C_{SW}^{14}$  of  $-43$  per mil and  $C_T^{14}$  of  $-160$  per mil, the predicted decrease in atmospheric  $C^{14}/C^{12}$  ratio would be about 1 per cent in 1954 compared with about 4 per cent based on the more complex model used in this paper. The reason for this difference is apparent.

The deep water masses act as infinite reservoirs for both radiocarbon and tritium. In one case more than adequate amounts of  $C^{14}$  can be returned to the atmosphere without altering the  $C^{14}/C^{12}$  ratio of the deep reservoirs, and in the other case all the tritium in the system could be added to the deep water reservoirs without creating a significant return flux to the surface.

A potential third test of the model would be to observe the distribution of bomb-produced

$C^{14}$  in the system. Again the test will be mainly on the characteristics of the surface water reservoirs. Whereas up to 1958 the magnitude of the bomb effect in the oceans was so small that precise experimental estimates were not obtained, Broecker and Olson [in press] have shown that the results obtained are consistent with the increase predicted by the model. Within a few years the effect will be of such magnitude that precise measurement will be possible.

Besides the tests provided by the distribution of tritium and the fossil  $CO_2$  effect, the case for the steady-state model is to some extent fortified because certain features of the  $C^{14}$  distribution itself can be predicted. The rather low  $\Delta C^{14}$  values obtained for Antarctic surface water samples are a natural result of the continual addition to this reservoir of low  $C^{14}$  water from the Atlantic and Pacific-Indian deep reservoirs to the vertically mixed Antarctic reservoir. Although more tenuous, the increase in  $\Delta C^{14}$  from the Antarctic to the South Atlantic to the Caribbean to the North Atlantic to the Arctic can be explained by progressive equilibration with the atmosphere as water is transported from south to north along the Atlantic surface. The Arctic reservoir could well show a reversal in this trend if a major portion of the water re-

TABLE 9. Distribution of  $C^{14}$  in the Ocean-Atmosphere System if the Oceanic Mixing Rates ( $R_i$ ) Were Double or Half the Computed Values ( $L$  is held constant.)

| Reservoir | Observed $\Delta C^{14}$ ,<br>‰ | Predicted $\Delta C^{14}$ if $R_i' = 2R_i$ ,<br>‰ | Predicted $\Delta C^{14}$ if $R_i' = \frac{1}{2}R_i$ ,<br>‰ |
|-----------|---------------------------------|---|---|
| 1         | 0                               | 0   | 0   |
| 2         | -32                             | -32   | -20   |
| 3         | -46 (-65)*                      | -66   | -49   |
| 4         | -35                             | -45   | -18   |
| 5         | -200                            | -145  | -280  |
| 6         | -113 (-120)*                    | -90   | -135  |
| 7         | -104                            | -85   | -153  |
| 8         | -28 (-30)*                      | -50   | -21   |

\* Estimates based on model.

turning to the surface from the Atlantic deep water system were returned to this reservoir instead of to the Antarctic. The higher Arctic value obtained by *Fonselius and Ostlund* [1959] (see Tables 2 and 4) thus lends support to the 'cycle' pattern for the Atlantic circulation as opposed to the 'exchange' pattern for the Pacific circulation.

Since with the exception of  $L$  no errors are attached to the estimates of the rates determined on the basis of the model, it is of interest to examine the sensitivity of these unknowns to the parameters selected. One way to approach this problem is to compare the steady-state  $\Delta C^{14}$  values predicted by the model if the three rates ( $R_i$ ) were twice (or half) the values given in Table 7 with the observed  $\Delta C^{14}$  values. This has been done in Table 9. The results have been normalized so that the atmospheric  $\Delta C^{14}$  value is zero in each case. The absolute  $C^{14}/C^{12}$  ratio in the atmosphere would be 7 per cent lower than it is at present in the first case and 8 per cent higher in the second. Inspection of these results indicates that if the rates were changed by 50 per cent, a significant deviation between the measured and computed  $\Delta C^{14}$  values for reservoirs 5, 6, and 7 would exist.

It is also necessary to consider the sensitivity of the model to changes in the thickness of the mixed layer. It is possible that the Central Water systems should at least in part be included in the mixed layer rather than entirely in the

deep reservoirs. In this case the effective thickness of the surface reservoirs could be as great as 300 meters. As  $L$  is independent of the oceanic mixing model selected, it remains unchanged. Examination of the steady-state equations shows that  $R_1$ ,  $R_2$ ,  $R_3$ ,  $x_4$ ,  $x_5$ , and  $x_6$  are also essentially independent of this depth. The mean residence times given in Figure 6 for the surface reservoirs (2, 3, and 4) become three times as great. This results in a change from the predicted value of 35 per mil for the industrial  $CO_2$  effect to a value of approximately 22 per mil. In the case of the tritium concentrations this would reduce the predicted values by a factor of slightly more than 2 for both the North Atlantic and Pacific surface water reservoirs. Thus increasing the depth of the surface reservoir brings the predicted and observed industrial effects closer together but drives the predicted and observed tritium concentrations farther apart.  $C^{14}$  data from future bomb tests should be very useful in this respect, as it will provide a third independent test.

Although the steady-state model given here incorporates only the gross features of large-scale circulation, it may prove useful in certain types of investigations. The rate of mixing of bomb-produced  $C^{14}$  with oceanic  $CO_2$ , the rate of oxygen consumption in the deep ocean, the uptake of industrial  $CO_2$  by the oceans, and the rate of interoceanic mixing are examples of problems which require a model of this type. Because of the restriction to problems involving time scales of more than 100 years, the natural radiocarbon method is far better suited to these large-scale problems than to detailed studies of mixing within a given local reservoir.

*Nonsteady-state models.* If mixing within the oceans occurs discontinuously on a time scale of tens to hundreds of years, results based on steady-state assumptions could be seriously in error. As an example a hypothetical case is demonstrated in Figure 8. It is assumed that the ocean mixes very rapidly during alternate 100-year periods. During the intervening 100-year periods no mixing occurs between the water above and below 100 meters in the ocean. The exchange rate between the atmosphere and the surface of the ocean is assumed to be unaltered by the factors causing the oceanic mixing rate to change. As shown in the figure, the atmosphere



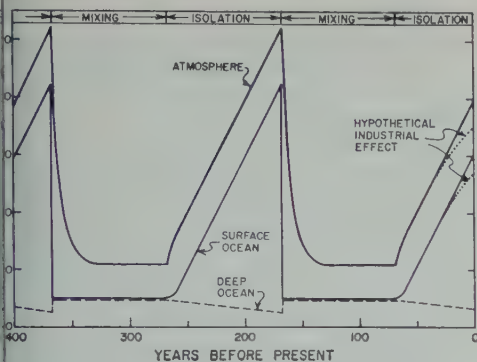


Fig. 8. A hypothetical example of the effects of steady-state mixing on the distribution of  $C^{14}$  in the ocean-atmosphere system. One-hundred-year periods of very rapid oceanic mixing are separated by 100-year periods during which no mixing occurs between the upper 100 meters of the ocean and the deep sea. The distribution of  $C^{14}$  at  $t = 0$  corresponds roughly to that of the present (corrected for industrial  $CO_2$  and bomb-produced  $C^{14}$ ).

Surface and deep oceans undergo large changes in  $\Delta C^{14}$  while the deep oceans show little change. The gradual rise in  $\Delta C^{14}$  during the isolation period is the result of storage of newly produced  $C^{14}$  in the relatively small atmospheric and surface ocean reservoirs. The sharp fall at the onset of mixing is the result of the transfer of this  $C^{14}$  into the much larger deep ocean reservoir. Note that the zero on the  $\Delta C^{14}$  scale is placed 25 per mil above the oceanic average corresponding to the present atmospheric value (corrected). Thus if the oceans were mixing in this hypothetical manner the present distribution would correspond to  $t = 0$  on the time scale of the figure.

The importance of this example is that it demonstrates the sensitivity of the atmospheric  $\Delta C^{14}$  value to changes in oceanic mixing rates. As pointed out by Broecker [1957] and by deVries [1958], measurement of the  $C^{14}/C^{12}$  ratio in tree rings of known growth date (providing estimates of the  $\Delta C^{14}$  value of the atmosphere at times in the past) should place limits on the extent of nonsteady-state mixing in the oceans.

Measurements available to date [deVries, 1958; Ostlund, 1957; Munnich and Vogel, 1958; Broecker, Olson, and Bird, 1959; Libby, Tauber, and Munnich, 1960] indicate that the atmospheric  $C^{14}/C^{12}$  ratio has

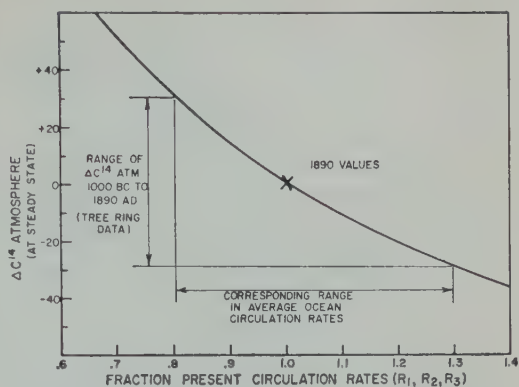


Fig. 9. Steady-state  $\Delta C^{14}$  values for atmospheric  $CO_2$  as a function of the transfer rate adopted for the oceanic mixing model. The three oceanic mixing rates are assumed to change by the same fraction. The maximum range of observed variations in the  $\Delta C^{14}$  value for the atmosphere over the last 1000 years is indicated by the vertical arrow and the corresponding maximum range of transfer rates by the horizontal arrow.

remained within  $\pm 25$  per mil of its 19th century value during the past 1000 years. There is a strong suggestion that real cyclic variations have occurred within these limits. Since part or even all of the observed variation could be the result of variations in the production rate of  $C^{14}$ , no definite conclusions regarding oceanic circulation variations can be made. It is possible, however, to place an upper limit on the amount of large-scale nonequilibrium circulation that has occurred over the past 1000 years. As shown in Figure 9, in order to produce the range of observed variations in atmospheric  $\Delta C^{14}$ , the rates  $R_1$ ,  $R_2$ , and  $R_3$  in the steady-state model would have to be varied only about 30 per cent on either side of the present values (assuming that the three rates undergo the same fractional change). It should be emphasized that, whereas the atmosphere is a sensitive indicator of large-scale variations in oceanic circulation rates, large changes in the rate for one local reservoir would not be detectable by this approach.

**Summary.** Although tree ring data suggest that the average turnover rate for the ocean has been relatively constant during the past 1000 years, it is not possible to state with any certainty whether any given local reservoir is at steady state. Because of the small volume



of the Atlantic Ocean compared with that of the Pacific and Indian Oceans it could well be considered a local reservoir. Thus it is entirely possible that part of the difference in the  $\Delta C^{14}$  values of surface and deep water in the Atlantic is the result of storage of newly produced  $C^{14}$  in the surface layers (Fig. 8) rather than to radioactive decay, as was assumed in the steady-state calculations. In this sense the steady-state estimates represent upper limits on the deep water residence times.

**Conclusions.** The results presented in this paper clearly demonstrate the value of the radiocarbon method in oceanographic studies. When used in conjunction with the proper oceanographic information and with other isotope data ( $H^3$ ,  $Si^{32}$ ,  $Ra^{226}$ ,  $Sr^{90}$ , . . .),  $C^{14}$  data will provide estimates of the large-scale mixing rates within the ocean. As many questions remain unanswered, the time estimates given here must be considered tentative. Perhaps the main value of the paper is that it (1) provides a broad base for more detailed future investigations, (2) demonstrates to some extent the potentialities as well as the limitations of the natural radiocarbon method in studies of oceanic circulation, and (3) points out some of the main problems requiring solution before more reliable time estimates can be made.

**Acknowledgments.** During the course of this investigation we have received aid from nearly all the members of the staff at Lamont. J. Ewing, W. Beckmann, and T. Takahashi were of particular help in collecting the samples. J. L. Kulp provided laboratory facilities and actively helped and encouraged the entire laboratory phase of the work. C. Tucek and E. A. Olson were responsible for the design and construction of much of the chemical system. M. Zickl, J. Hubbard, and S. Friedman were of considerable aid in carrying out many of the technical operations involved, and F. Gwiner, W. Tamminga, and A. Ludas were responsible for constructing the shipboard sampling and processing apparatus. We are extremely grateful for this aid. Without it the present research could not have been accomplished. Financial support was provided by the Atomic Energy Commission (grant AT(30-1)1808) from 1955 to 1959 and by the International Geophysical Year (grant Y/9.11/134).

#### REFERENCES

- Arnold, J. R., and E. C. Anderson, The distribution of  $C^{14}$  in nature, *Tellus*, **9**, 28-32, 1957.
- Begamann, F., and W. F. Libby, Continental water balance, ground-water inventory, and storage times, surface ocean mixing rates and worldwide water circulation patterns from cosmic and bomb tritium, *Geochim. et Cosmochim. Acta*, **12**, 277-296, 1957.
- Bolin, B., On the use of tritium as a tracer in water in nature, *Second United Nations Int. Conf. on the Peaceful Uses of Atomic Energy*, **18**, 336-343, 1958.
- Bolin, B., and E. Eriksson, Changes in the carbon dioxide content of the atmosphere and sea due to fossil fuel combustion, *Rosby Memorial Volume*, Rockefeller Inst. Press, New York, **13**, 142, 1959.
- Bowen, V. T., and L. T. Sugihara, Strontium-90 in the "mixed layer" of the Atlantic Ocean, *Nature*, **186**, 71-72, 1960.
- Brannon, H. R., A. C. Daughtry, D. Perry, W. Whitaker, and M. Williams, Radiocarbon evidence on the dilution of atmospheric and oceanic carbon by carbon from fossil fuels, *Trans. Am. Geophys. Union*, **38**, 643-650, 1957.
- Brodie, J. W., and R. W. Burling, Age determinations of southern ocean waters, *Nature*, **1**, 107-108, 1958.
- Broecker, W. S., Application of radiocarbon to oceanography and climate chronology, Ph.D. Thesis, Columbia University, 1957.
- Broecker, W. S., W. M. Ewing, R. Gerard, B. Heezen, and J. L. Kulp, The significance of variations of light isotope abundances in oceanographic studies, *Natl. Acad. Sci. Natl. Research Council Publ.* **572**, 118-134, 1958.
- Broecker, W. S., and E. A. Olson, Lamont radiocarbon measurements, **6**, *Am. J. Sci. Radiocarbon Suppl.*, **1**, 111-132, 1959.
- Broecker, W. S., and E. A. Olson, Further information on radiocarbon from nuclear tests, *Science*, in press.
- Broecker, W. S., E. A. Olson, and J. Bird, Radiocarbon measurements on samples of known age, *Nature*, **183**, 1582-1584, 1959.
- Broecker, W. S., C. S. Tucek, and E. A. Olson, Radiocarbon analysis of oceanic  $CO_2$ , *Appl. Radiation and Isotopes*, **7**, 1-18, 1959.
- Broecker, W. S., and A. Walton, Radiocarbon from nuclear tests, *Science*, **130**, 309-314, 1959.
- Broecker, W. S., and A. Walton, The geochemistry of  $C^{14}$  in fresh-water systems, *Geochim. et Cosmochim. Acta*, **16**, 15-38, 1959b.
- Burling, R. W., and D. M. Garner, A section in  $C^{14}$  activities of sea water between 9°S and 66°S in the southwest Pacific Ocean, *New Zealand Geol. Geophys.*, in press.
- Craig, H., The geochemistry of the stable carbon isotopes, *Geochim. et Cosmochim. Acta*, **1**, 53-92, 1953.
- Craig, H., The natural distribution of radiocarbon and the exchange time of carbon dioxide between atmosphere and sea, *Tellus*, **9**, 1-17, 1957.
- Craig, H., Distribution, production rate, and possible solar origin of natural tritium, *Phys. Rev.* **105**, 1125-1127, 1957b.

- ies, H., Variation in concentration of radio-carbon with time and location on earth, *Koninkl. d. Akad. Wetenschap. Proc., Ser. B*, 61, 1-9, 1958.
- usson, G. J., Reduction of atmospheric radio-carbon concentration by fossil fuel carbon di-oxide, *Proc. Roy. Soc. London, A*, 243, 561-574, 1958.
- elius, S., and G. Ostlund, Natural radiocarbon measurements on surface water from the North Atlantic and the Arctic Sea, *Tellus*, 11, 77-82, 1959.
- er, D. M., A radiocarbon profile in the Tas-man Sea, *Nature*, 182, 466-468, 1958.
- ti, B. J., F. Bazan, and J. L. Kulp, The geo-chemistry of tritium, *Trans. Am. Geophys. Union*, 39, 807-818, 1958.
- es, F. N., E. C. Anderson, and J. R. Arnold, Liquid scintillation counting of natural radio-carbon, *Intern. Conf. on the Peaceful Uses of Atomic Energy*, 14, 188-192, 1955.
- o, J. L., L. E. Tryon, W. R. Eckelmann, and A. Snell, Lamont natural radiocarbon mea-surements, 2, *Science*, 116, 409-414, 1952.
- nnich, K. O., and J. C. Vogel, Radioactive car-bon in the atmosphere produced by atomic ex-plosions, *Naturwissenschaften*, 14, 327-329, 1958.
- and, H. G., Stockholm natural radiocarbon measurements, 1, *Science*, 126, 493-497, 1957.
- er, T. A.,  $C^{14}$  variations in nature and the effect on radiocarbon dating, *New Zealand J. Technol.*, B, 37, 20-38, 1955.
- er, T. A., and G. J. Fergusson, Atmospheric radiocarbon as a tracer in geophysical circula-tion problems, *Second United Nations Intern. Conf. on the Peaceful Uses of Atomic Energy*, 526-532, 1958.
- n, M., and C. Alexander, U. S. Geological survey radiocarbon dates, 4, *Science*, 127, 1476-1477, 1958.
- mel, H., A survey of ocean current theory, *Deep-Sea Research*, 4, 149-183, 1957a.
- Stommel, H., Toward a future study of the abyssal circulation, Symposium on abyssal circulation, IUGG, Toronto Meeting, 1957b.
- Stommel, H., The abyssal circulation, *Deep-Sea Research*, 5, 80-82, 1958.
- Suess, H. E., Radiocarbon concentration in mod-ern wood, *Science*, 122, 415-417, 1955.
- Suess, H. E., N. W. Rakestraw, and H. Oeschger, Apparent age of deep water in the Pacific Ocean, *Preprints, Intern. Oceanog. Congr.*, p. 440, 1959.
- Suess, H. E., and R. Revelle, Carbon dioxide ex-change between atmosphere and ocean and the question of an increase of atmospheric  $CO_2$  dur-ing the past decades, *Tellus*, 9, 18-27, 1957.
- Sverdrup, H. U., Oceanography, chap. 5, in *The Earth as a Planet*, edited by G. P. Kuiper, Univ. Chicago Press, 1954.
- Sverdrup, H. U., M. W. Johnson, and R. H. Flem-ing, *The Oceans*, Prentice-Hall, Inc., New York, 1942.
- Swallow, C., and L. V. Worthington, Measure-ments of deep currents in the western North Atlantic, *Nature*, 179, 1183-1184, 1957.
- Willis, E. H., H. Tauber, and K. O. Munnich, Variations in the atmospheric radiocarbon con-centration over the past 1300 years, *Am. J. Sci. Radiocarbon Suppl.*, 2, 1-4, 1960.
- Wüst, G., Die Stratosphäre, Deutsche Atlantische Exped. *Meteor* 1925-1927, *Wiss. Erg.*, Bd. 6, Teil, 2, Lief., 288 pp., 1935.
- Wüst, G., Quantitative Untersuchungen zur Statik und Dynamik des Atlantischen Ozeans, Deutsche Atlantische Exped. *Meteor* 1925-1927, *Wiss. Erg.*, Bd. 6, 2 Teil, 6. Lief., 261-420, 1957a.
- Wüst, G., On the meridional volume transport within different strata of the South Atlantic, especially in abyssal depths, Symposium on Abyssal Circulation, IUGG, Toronto Meeting, 1957b.

(Manuscript received April 21, 1960; revised May 25, 1960.)



# Volume Diffusion as a Mechanism for Discordant Lead Ages

G. R. TILTON

*Geophysical Laboratory, Carnegie Institution of Washington  
Washington 8, D. C.*

**Abstract.** The discordant U-Pb ages of sixteen minerals having  $Pb^{207}/Pb^{206}$  ages of 2400 to 2800 m.y. closely fit a 2800 to 600 m.y. chord when placed on a  $Pb^{207}/U^{235}$ - $Pb^{206}/U^{238}$  diagram. These minerals are from the continents of North America, Africa, Europe, Australia, and Asia. If the data are interpreted as evidence of an episode of loss of lead 600 m.y. ago from minerals that were crystallized 2800 m.y. ago, it is strange that the same time of loss is indicated for all four continents. Moreover, most of the samples are from shield areas where no evidence has been found for metamorphic events 600 m.y. ago.

As an alternative explanation, lead may be considered to diffuse continuously from crystals at a rate governed by a diffusion coefficient ( $D$ ), the effective radius ( $a$ ), and the concentration gradient. Calculations of the present-day  $Pb^{206}/U^{238}$  and  $Pb^{207}/U^{235}$  ratios as a function of the parameter  $D/a^2$  yield ratios that lie on a 2800 to 600 m.y. chord on a  $Pb^{206}/U^{238}$ - $Pb^{207}/U^{235}$  diagram for losses of up to two-thirds of the lead produced in the sample.

This method is capable of explaining the regularity of the age patterns from one continent to another and does away with the need for episodic losses of lead simultaneously in all the areas 600 m.y. ago. Other examples have been found of suites of minerals with discordant ages which fit this hypothesis; they include minerals with apparent ages of 1900, 1700, and 1100 m.y. The behavior of uranium-lead ages in several zircons from metamorphic rocks is indicative of an activation energy for diffusion of lead of less than 10 kcal/mol over a temperature range of 50° to 400°-600°C. This inference, when considered with available data for the diffusion of argon from micas, is capable of explaining why uranium-lead ages measured from zircon may be either greater or less than the potassium-argon ages of coexisting mica.

## INTRODUCTION

When the first isotopic lead ages were published by Nier [1939], a problem has existed concerning lack of agreement between the  $Pb^{206}$ - $U^{238}$  and  $Pb^{207}$ - $U^{235}$  ages obtained on many mineral samples. In most, if not all, cases it can be assumed that minerals having discordant uranium-lead ages have not been closed systems with respect to uranium or lead or both since they were first formed. The cause of the age discrepancies is a matter of some importance since, in principle, such ages contain information about the post-crystallization history of a mineral that cannot be obtained from concordant lead ages. The present paper deals with the problem of interpreting discordant ages found in minerals from rocks taken largely from Precambrian shield areas.

Certain regularities in the discordant uranium-lead ages of monazites from Rhodesia and Madagascar and of uraninite and thucholite from the Witwatersrand were discussed by Ahrens [1955a, 1956]. Ahrens devised several methods of plotting

the ages, and he believed that the regularities in the discordant age patterns suggested 'control of lead loss by physical processes which perhaps have operated at varying degrees but at a uniform rate ever since the minerals were formed.' Wetherill [1956a, 1956b] then showed that the regularities noted by Ahrens could be explained by a simple geochemical process. For example, the Rhodesian age pattern could be produced by minerals with an age of 2700 m.y. that lost different fractions of their lead 500 m.y. ago. Wetherill developed a graphical procedure that permits rapid determination of the age and time of lead removal for suites of cogenetic minerals that have experienced a single episode of lead loss. The procedure is as follows: on a plot of  $Pb^{207}/U^{235}$  against  $Pb^{206}/U^{238}$ , a curve 'concordia,' is drawn representing the locus of all points such that  $t(Pb^{207}/U^{235}) = t(Pb^{206}/U^{238})$ . Samples having an age  $t_1$  that lost lead at  $t_2$  will lie on a chord connecting  $t_1$  and  $t_2$ , their position along the chord depending on the amount of lead that was lost.



*Russell and Ahrens* [1957] studied four suites of minerals with ages from 1000 to 2700 m.y. They noted that, if a single episode of removal of lead is assumed to be the cause of the age discordances, the time of lead removal appears to be a function of the age of a given suite of minerals. The younger a suite of minerals, the more recent the time of loss of lead appears to be. The authors suggested that a physical rather than a chemical process seems to be required to explain this regularity. The process should be one that causes a greater percentage loss of  $Pb^{207}$  than of  $Pb^{206}$ . Russell and Ahrens postulated that the differential losses might be related to differences in recoil energies from  $\alpha$ -particle emission of various members of the two uranium decay series or to some 'wholly unforeseen process.' Serious questions can be raised about the samples used by Russell and Ahrens. Of four mineral suites, two are from the Goldfields District in Saskatchewan and one is from the Witwatersrand. As was mentioned by the authors, it is recognized that the Goldfields pitchblendes represent more than one generation. The isotopic data indicate more than one episode of lead

loss—one about 1200 and one about 200 m.y. ago [*Eckelmann and Kulp*, 1956; *Aldrich and Wetherill*, 1958]. The episode of loss at 1200 m.y. does not fit the authors' proposed regularity. For the Witwatersrand samples, the spread of points on a  $Pb^{207}/U^{235}-Pb^{206}/U^{238}$  diagram [*Wetherill*, 1956a] indicates a complex history that involves more than one age of sample or time of lead loss, or both. The test of possible simple regularities in discordant lead ages should be limited to areas having simpler geological histories.

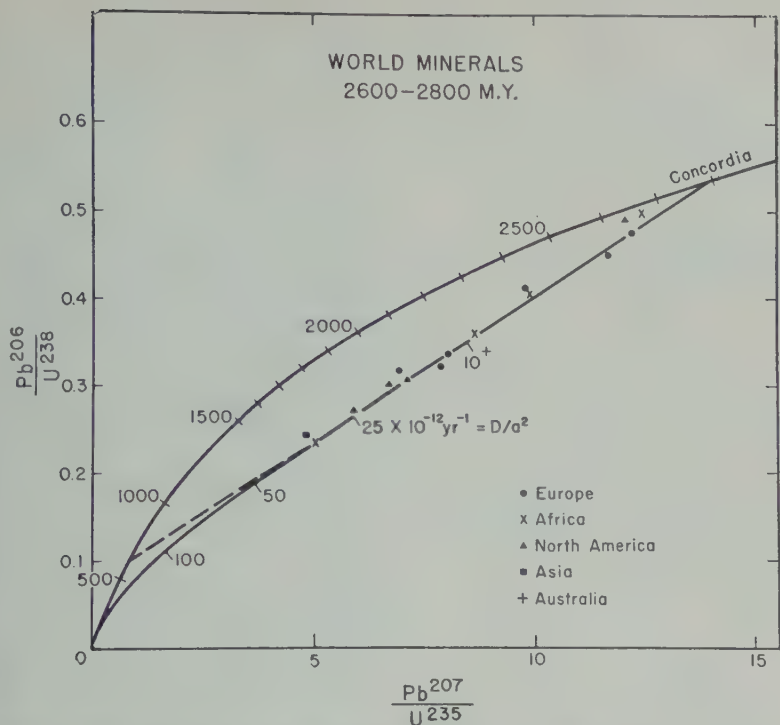
Since these earlier papers were written, the number of minerals having  $Pb^{207}-Pb^{206}$  ages about 2600 m.y. has been considerably increased. Currently available results are given in Table 1 which includes only those minerals having  $Pb^{207}-Pb^{206}$  ages  $\geq 2300$  m.y. Minerals with still lower  $Pb^{207}-Pb^{206}$  ages that appear to belong to this group have been omitted to avoid possible erroneous inclusion of younger minerals.

If the samples from Table 1 are plotted on a  $Pb^{207}/U^{235}-Pb^{206}/U^{238}$  diagram, an array of points is obtained that corresponds to the case of a group of minerals having ages of 2600 to 2800

TABLE 1. Minerals Having  $Pb^{207}-Pb^{206}$  Ages of 2300–2800 m.y.\*

| Location                     | Mineral      | Age, million years         |                            |                             |                             | Reference                     |
|------------------------------|--------------|----------------------------|----------------------------|-----------------------------|-----------------------------|-------------------------------|
|                              |              | $\frac{Pb^{206}}{U^{238}}$ | $\frac{Pb^{207}}{U^{235}}$ | $\frac{Pb^{207}}{Pb^{206}}$ | $\frac{Pb^{208}}{Th^{232}}$ |                               |
| Mätäsvaara, Finland          | Zircon       | 2520                       | 2650                       | 2770                        | 2750                        | Tilton and Kouvo, unpublished |
| Heinävaara, Finland          | Zircon       | 2250                       | 2240                       | 2620                        | 1280                        | Tilton and Kouvo, unpublished |
| Koli, Finland                | Zircon       | 1890                       | 2270                       | 2650                        | 1790                        | Tilton and Kouvo, unpublished |
| Huhtilampi, Finland          | Zircon       | 1820                       | 2240                       | 2640                        | 1820                        | Tilton and Kouvo, unpublished |
| Sotkuma, Finland             | Zircon       | 1810                       | 2150                       | 2470                        | 1760                        | Tilton and Kouvo, unpublished |
| Saksagan, Repikhovo, Ukraine | Orthite      | 2420                       | 2610                       | 2780                        | 2630                        | Vinogradov, 1956              |
| Huron Claim, Manitoba        | Uraninite    | 1560                       | 1980                       | 2480                        | 1280                        | Nier, 1939                    |
| Dickinson Co., Michigan      | Zircon       | 1740                       | 2150                       | 2570                        | 1250                        | G. L. Davis, unpublished      |
| Dickinson Co., Michigan      | Zircon       | 1710                       | 2100                       | 2510                        | 1300                        | G. L. Davis, unpublished      |
| Cooke City, Montana          | Uraninite    | 2600                       | 2650                       | 2700                        | ...                         | Gast and others, 1958         |
| Ebonite Claims, S. Rhodesia  | Monazite     | 2640                       | 2670                       | 2700                        | 2640                        | Holmes, 1954                  |
| Jack Tin Claims, S. Rhodesia | Monazite     | 2210                       | 2460                       | 2660                        | 1940                        | Holmes, 1954                  |
| Irumi Hills, N. Rhodesia     | Monazite     | 1990                       | 2330                       | 2640                        | 1380                        | Holmes, 1954                  |
| Antsirabi, Madagascar        | Monazite     | 1370                       | 1850                       | 2450                        | 610                         | Holmes, 1954                  |
| Yadiur, India                | Monazite     | 1410                       | 1820                       | 2330                        | 1800                        | Holmes, 1955                  |
| Woodstock, W. Australia      | Tanteuxenite | 1900                       | 2360                       | 2790                        | 2590                        | Greenhalgh and Jeffery, 1959  |

\* Minerals having  $Pb^{206}-U^{238}$  ages that are greater than their  $Pb^{207}-Pb^{206}$  ages have been excluded from this table for reasons stated in the text.



1. Parent-daughter ratios for minerals having  $\text{Pb}^{207}\text{-Pb}^{200}$  ages of 2300 to 2800 m.y. The curve through the points is calculated for loss of lead by continuous diffusion from spheres of radius  $a$  having ages of 2800 m.y.

that lost varying amounts of lead with time to uranium 500 to 600 m.y. ago (see Table 1). If the discordances are to be explained by a single episode of lead loss, the time of loss would be about the same from one continent to another and negligible losses occurred from 2700 to 2800 m.y. ago. This would seem to be a rather remarkable coincidence.

Further difficulty arises from the fact that South African and European minerals are from shield areas where age-determination is not, as yet, given evidence for radiogenic activity or metamorphism 500 to 600 m.y. ago. Minerals having ages in this interval occur in South Africa [Holmes and Cahen, 1954] but it is not known to what extent the  $\text{Pb}/\text{U}$  ratios in the Rhodesian minerals in Table 1 were affected by this episode of discordant crystallization.

#### VOLUME DIFFUSION

*Mathematical treatment.* If ages of 2600 to 2800 m.y. are accepted for the samples shown

in Table 1, any mechanism postulated to explain the discordant ages must be one that produces a greater loss of radiogenic  $\text{Pb}^{207}$  than of  $\text{Pb}^{200}$ . Episodic loss of lead at some time in the past is one such mechanism. Diffusion also satisfies this requirement in at least a qualitative way. Owing to the shorter half-life of  $\text{U}^{235}$ ,  $\text{Pb}^{207}$  has a greater mean age than the  $\text{Pb}^{200}$  in a mineral and thus has had a longer time to diffuse out.

A quantitative solution for the loss of a radiogenic daughter product by diffusion can be obtained for certain idealized cases. The assumptions used here are:

1. The mineral crystals are spheres of effective radius  $a$ . The value of  $a$  may be determined by grain boundaries or by lattice imperfections and need not be related directly to the physical size or shape of the grains.
2. Uranium is distributed uniformly within the spheres.
3. Diffusion of uranium and intermediate daughter products is negligible compared with that of lead.

4. The diffusion coefficient,  $D$ , is constant over the time the mineral has existed.

5. The diffusion of lead is governed by Fick's law, so that the differential equation for the change of lead concentration with time for any radial volume element in the sphere may be written

$$\frac{\partial C}{\partial t} = D \left( \frac{\partial^2 C}{\partial r^2} + \frac{2}{r} \frac{\partial C}{\partial r} \right) + \lambda N_0 e^{-\lambda t'} \quad (1)$$

where

$C$  = atom concentration of lead daughter product.

$N_0$  = initial concentration of uranium parent atoms.

$\lambda$  = decay constant of the uranium parent.

$t'$  = time measured from the beginning of the diffusion process.

The boundary conditions are:

$$C = 0 \quad t = 0$$

$$C = 0 \quad \text{all } t \quad r = a$$

A solution for equation 1 can be derived in a simple way from the equation for the diffusion of a component out of a sphere, for uniform initial distribution [Barrer, 1951, pp. 28-29]:

$$\frac{\bar{C}}{C_0} = \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{e^{-n^2 \pi^2 D t / a^2}}{n^2} \quad (2)$$

where  $C_0$  is the initial concentration and  $\bar{C}$  is the average concentration after an elapsed time  $t$ . By radioactive decay uranium has continually generated lead so that, at time  $t'$ ,

$$C_0(t') = \lambda N_0 e^{-\lambda t'} \quad (3)$$

The average concentration of lead at the present time as the result of such a process is

$$\bar{C} = \frac{6}{\pi^2} \int_0^t \sum_{n=1}^{\infty} \frac{\lambda N_0 e^{-\lambda t' - n^2 \pi^2 D (t-t') / a^2}}{n^2} dt' \quad (4)$$

where  $t$  is the age of the mineral.

Equation 4 integrates to

$$\frac{\bar{C}}{N} = \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{\lambda (e^{\lambda t - n^2 \pi^2 D t / a^2} - 1)}{n^2 (\lambda - n^2 \pi^2 D / a^2)} \quad (5)$$

where  $N$  is the concentration of uranium (atoms) at present.

Equation 5 is equivalent to the solution given

first by Wasserburg [1954]. Nicolaysen [1954] published an extensive table of values for equation 5. His tables were used in preparing the figures in the present paper.

*Solutions of the diffusion equation for a sphere.* Figure 1 shows a quantitative comparison between the discordant age pattern expected for hypothetical 2800 m.y.-old minerals that have lost lead by the diffusion model given above and the discordant ages of the minerals in Table 1. The curve labeled with values of  $D/a^2$  is the locus of points which represent possible samples that conform to the diffusion model. For example, the point at  $10 \times 10^{-12}$  yr<sup>2</sup> gives the  $\text{Pb}^{207}/\text{U}^{235}$  and  $\text{Pb}^{206}/\text{U}^{238}$  values that exist today in a mineral having an age of 2800 m.y. and a value of  $10 \times 10^{-12}$  for  $D/a^2$  over the past 2800 m.y. For values of  $D/a^2$  up to approximately  $50 \times 10^{-12}$  yr<sup>-1</sup>, diffusion loss would produce a group of samples that closely fit the 2800 to 600 m.y. chord. A group of minerals having an age of 2800 m.y. that lost different proportions of lead 600 m.y. ago, this being the only time at which any loss of lead occurred, would fit the same chord. Thus a process of episodic loss of lead that took place continuously since the time these minerals crystallized can produce the same pattern of discordant ages as a process involving an episodic loss of lead. Because of this, the time of apparent episodic loss of lead found in Figure 1 need not have any particular geologic significance.

This explanation offers a solution to the problems that arise from an interpretation based on episodic lead loss. The discordant ages in Finland and Michigan can be explained without resort to a metamorphic episode 500 to 600 m.y. ago, for which no evidence is present. As mentioned above, minerals with ages of 450 to 600 m.y. do occur in southern Africa [Holm and Cahen, 1957], so that episodic loss is a possible explanation for the discordant age pattern of the African monazites. The diffusion mechanism for loss of lead helps to explain the linear relationship of samples from one continent to another in Figure 1. The diffusion model requires only that the samples be of approximately the same age, whereas the episodic-loss mechanism requires all the samples to be of the same age and all to have lost lead 500 to 600 m.y. ago.

The  $D/a^2$  curve ('discordia') shown in Figure



TABLE 2. Uranium-Lead Ages from Belomorja, Karelia

| Mineral    | Age, million years                       |  |  | Reference    |
|------------|--|--|--|--------------|
|            | $\frac{\text{Pb}^{206}}{\text{U}^{238}}$ | $\frac{\text{Pb}^{207}}{\text{U}^{235}}$ | $\frac{\text{Pb}^{207}}{\text{U}^{206}}$ |              |
| Uraninite  | 1760                                     | 1800                                     | 1870                                     | Starik, 1956 |
| Uraninite  | 1660                                     | 1760                                     | 1880                                     | Starik, 1956 |
| Uraninite  | 1680                                     | 1740                                     | 1820                                     | Starik, 1956 |
| Uraninite  | 1050                                     | 1300                                     | 1800                                     | Starik, 1956 |
| Uraninite  | 1150                                     | 1380                                     | 1780                                     | Starik, 1956 |
| Monazite   | 1250                                     | 1450                                     | 1800                                     | Zhirov, 1958 |
| Monazite   | 1220                                     | 1440                                     | 1760                                     | Zhirov, 1958 |
| Thucholite | 1480                                     | 1655                                     | 1900                                     | Zhirov, 1958 |
| Thucholite | 1250                                     | 1460                                     | 1850                                     | Zhirov, 1958 |
| Thucholite | 1120                                     | 1340                                     | 1770                                     | Zhirov, 1958 |
| Thucholite | 850                                      | 1170                                     | 1810                                     | Zhirov, 1958 |

tended only as an example of a solution of the diffusion mechanism, not as a least-squares fit of the data. Since the samples are likely to be of exactly one age, any such fit would be meaningless.

Playesen [1957] published values for the ages of a number of samples after correcting for discordant ages for loss of lead by diffusion. His graphical method of solution differs from the method used here but leads to the same solution for the parent age. He showed that the South African Yadiur monazites have ages of 2600 m.y. if diffusion according to the model described above is responsible for the discordant ages. His method does not show the lead-loss relationships illustrated by Figure 1.

A second example of a suite of minerals with discordant ages that fit the diffusion hypothesis is given in Table 2 and Figure 2. These minerals come from pegmatites in the White Sea region of Karelia. The fit of points to the  $D/a^2$  curve in Figure 2 is substantially better than in Figure 1. This might be expected since the requirement that the minerals all have the same age is more likely to be true for the Karelian suite than for samples taken from various continents. In preparation of Figure 2, several monazites having  $\text{U}^{238}$ - $\text{Pb}^{206}$  ages greater than their  $\text{Pb}^{207}$ - $\text{Pb}^{208}$  ages were rejected since such samples obviously do not conform to the basic assumptions of the diffusion calculations. The thucholite ages are approximately equivalent to uraninite ages since

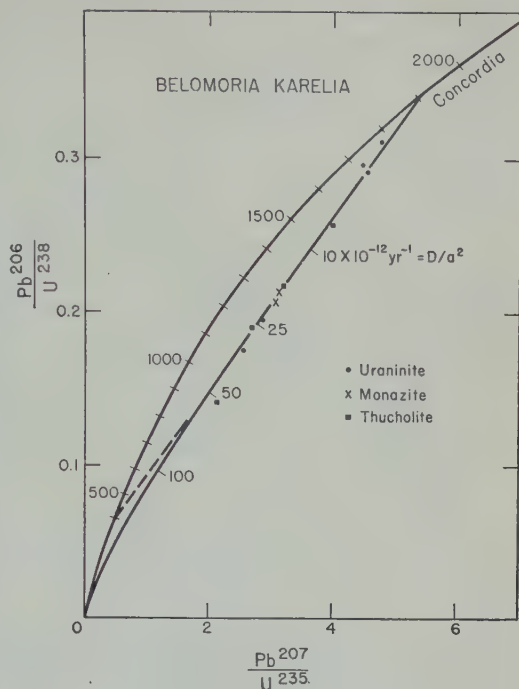


Fig. 2. Parent-daughter ratios for minerals from Belomorja, Karelia, compared with calculated ratios for loss of lead by continuous diffusion.

thucholites from Finland have been found to be uraninite embedded in an organic matrix (Kouvo, private communication).

Figure 3 compares the U-Pb ratios of a number of North American minerals having  $\text{Pb}^{207}$ - $\text{Pb}^{208}$  ages of 1000 to 1150 m.y. against a  $D/a^2$  curve for 1150-m.y.-old minerals. For minerals this young, 'discordia' lies closer to 'concordia' than in the two previous examples, and the time of apparent lead loss is closer to 0 m.y. so that interpretation of the data becomes more difficult. Clearly, however, the diffusion hypothesis is capable of explaining the discordant ages. This explanation has some merit because the majority of minerals in Figure 3 have  $\text{U}^{238}$ - $\text{Pb}^{206}$  ages about 10 per cent lower than their  $\text{Pb}^{207}$ - $\text{Pb}^{208}$  ages. If episodic loss of lead were the cause of the discordance, a more erratic spread of ages might be expected. In an earlier paper dealing with the isotopic zircon ages [Tilton, Davis, Wetherill, and Aldrich, 1957], the repeated finding of  $\text{U}^{238}$ - $\text{Pb}^{206}$  ages that were 10 per cent lower than 207-206 ages led the authors to consider



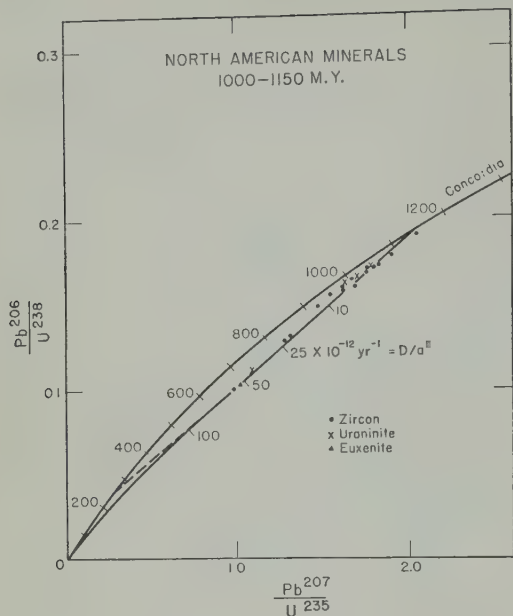


Fig. 3. Parent-daughter ratios for North American minerals with  $\text{Pb}^{207}$ - $\text{Pb}^{206}$  ages of 1000 to 1150 m.y. The curve through the points is calculated for loss of lead by continuous diffusion.

the possibility that this difference might be due to loss of lead by faulty analytical procedures. Subsequent work has shown the analytical procedures to be reliable and age discordances real.

In western North Carolina and eastern Tennessee three zircons have discordant ages that do not fit the diffusion-loss hypothesis. These samples have age discordances that might be the result of an episodic loss of lead about 450 m.y. ago superimposed on diffusion losses. These zircons have been omitted from Figure 3.

A final diffusion comparison is made in Figure 4. These are zircon age determinations taken from the work of Kuovo [1958]. The line of best fit to the five points passes through the origin, but, until more analyses are obtained, a diffusion mechanism for lead loss cannot be excluded. The zircons are from a shield area where there is no evidence for metamorphism in recent time.

**Effect of temperature changes.** The diffusion calculations contain the implicit assumption, which would seem to be unrealistic, that temperature has remained constant over periods of billions of years. The effect of changes in tem-

perature on losses of lead by diffusion is difficult to evaluate. Equation 1 can be solved analytically for two cases: (a) no lead is present initially; (b) any such lead is homogeneously distributed. Calculations for temperature changes involve changing  $D$  after concentration gradients have already been established for lead as a result of previous diffusion. No analytical solutions exist for this problem.

Some information about the effect of temperature changes can be obtained from application of equation 5. Consider the case of a mineral with an age of 2800 m.y. which lost lead as a result of an increase in the diffusion rate for a brief interval of time, e.g., 10 to 100 m.y., 11 m.y. ago. At all other times let lead be lost at a rate governed by a particular value for  $D$ , say  $10 \times 10^{-12} \text{ yr}^{-1}$ . The values 2800 and 11 m.y. are chosen because the rubidium-strontium ages of biotite associated with several of Finnish zircons in Table 1 are 1800 m.y. Tilton with other zircon work indicates that 1800 to 1900 m.y. ago was a time of metamorphism in the area (Wetherill, Kuovo, Tilton, and Cresswell, paper to be submitted to the *Journal of Geology*). The concentrations of  $\text{Pb}^{206}$  and  $\text{Pb}^{207}$  which existed 1800 m.y. ago in a mineral which has an age of 2800 m.y. can be calculated from equation 5. Of the lead present at that time, some is to be lost by heating, the loss being governed by Fick's law with the concentration of lead at zero at the boundary of the crystal, then more lead is to be lost at a rate governed by  $D/a^2$ , say  $10 \times 10^{-12} \text{ yr}^{-1}$  from 1800 m.y. to the present.

Two limiting cases are obvious: as the loss of lead at 1800 m.y. approaches zero, the  $\text{Pb}^{207}/\text{U}^{235}$  and  $\text{Pb}^{206}/\text{U}^{238}$  ratios approach the point 10 on the  $D/a^2$  curve for a 2800-m.y.-old mineral in Figure 1. As lead loss approaches 100 per cent, the ratios approach the point 10 on the  $D/a^2$  curve for an 1800-m.y.-old mineral. For intermediate cases, the problem reduces to following the change in isotopic composition of lead as a function of the amount of lead lost by diffusion. The amount of lead lost by diffusion from 1800 m.y. ago has been generated in a mineral for 1000 m.y. This can be done by calculating from equation 5 the  $\text{Pb}^{207}$  and  $\text{Pb}^{206}$  at 1800 m.y. and the amounts at successive times such as 800, 600, 400, 200 m.y. ago are calculated. The radiogenic lead formed from 1800 m.y. ago to the time of consideration can be calculated and subtracted from the total lead present at 1800 m.y. ago to give the lead present at the time of consideration.

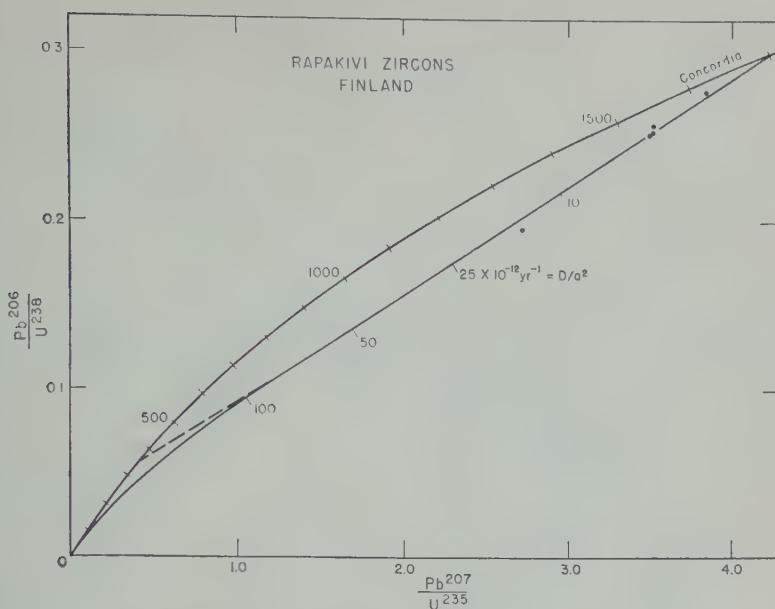


Fig. 4. Parent-daughter ratios for zircon from Finnish rapakivi granites compared with calculated ratios for loss of lead by continuous diffusion.

from the total lead present at that time. The reference represents what remains of the lead that was present 1800 m.y. ago. As the loss of lead approaches 100 per cent, the isotopic composition of the residual lead approaches that of lead generated from 2800 to 1800 m.y. ago in a mineral that experienced no loss of lead, as can be seen from equation 2. For a very long elapsed time after generation of the lead, the fraction of lead lost will be almost independent of the time interval in which it was generated. The  $Pb^{207}/Pb^{206}$  ratios 1800 m.y. ago in minerals that are now 2800 m.y. old were 0.3122 for  $D/a^2 = 0$ , 0.3061 for  $D/a^2 = 10 \times 10^{-12} \text{ yr}^{-1}$ , 0.3031 for  $D/a^2 = 20 \times 10^{-12} \text{ yr}^{-1}$ , and 0.3005 for  $D/a^2 = 30 \times 10^{-12} \text{ yr}^{-1}$ .

These calculations show that the  $Pb^{207}/Pb^{206}$  ratios of the last three cases change in a regular manner toward 0.3122 as lead is lost by further diffusion. Since the ratios do not differ greatly from 0.3122 to start with, it can be assumed as an approximation that  $Pb^{207}$  and  $Pb^{206}$  are lost in amounts that are essentially proportional to their isotopic abundances at 1800 m.y. The resulting array of discordant ages

can be plotted as in Figure 5. Loss of half of the lead contained in the minerals at 1800 m.y. would yield points that plot approximately, but not exactly, at the mid-points of the lines connecting the  $D/a^2$  curves. (The points will actually be displaced toward the 2800-m.y. curve.) Loss of 20 to 30 per cent of lead 1800 m.y. ago would not produce a serious departure from the  $D/a^2$  curve for 2800-m.y.-old minerals. Such losses are quite within the spread noted for the five Finnish zircons, for example. Any episodic loss of lead 2500 to 2800 or 0 to 1000 m.y. ago would be even more difficult to detect.

The method illustrated in Figure 5 can be used to predict trends resulting from changes in the value of the diffusion coefficient, but it does not give information on the amount of lead existing in a mineral at a particular time that will be lost for a given change in  $D$ . More elaborate procedures for solving the diffusion equation after concentration gradients have been established by diffusion are required.

*Diffusion from crystals with nonspherical shapes.* The foregoing discussions have dealt with diffusion losses from crystals of spherical shape. Since this is an idealization not realized in nature, it is worth while to examine the effect of changes in crystal shape. The calculations for

Decay constants used:  $U^{238}$ :  $1.54 \times 10^{-10} \text{ yr}^{-1}$ ;  $U^{235}$ :  $9.72 \times 10^{-10} \text{ yr}^{-1}$ .

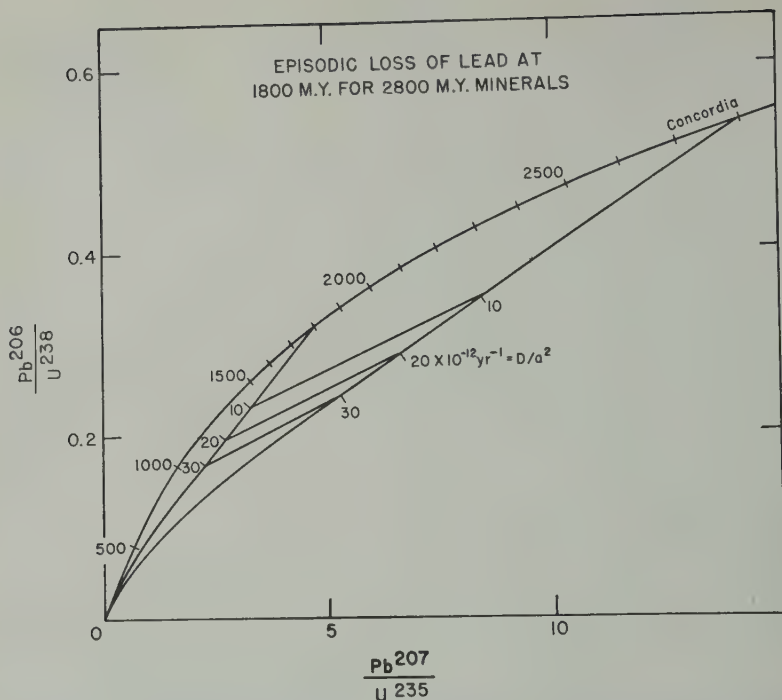


Fig. 5. Parent-daughter ratios resulting from episodic loss of lead 1800 m.y. ago superimposed on continuous loss of lead by diffusion for hypothetical minerals having an age of 2800 m.y.

the sphere can also be made for the infinite cylinder, that is, a cylinder in which diffusion takes place in the radial direction only. If the same assumptions outlined above for the sphere are used for the cylinder an equation for the average lead content of a cylinder of radius  $a$  may be derived by the same method used for the sphere, obtaining

$$\frac{\bar{C}}{N} = 4 \sum_{n=1}^{\infty} \frac{\lambda(e^{\lambda t - \xi_n^2 D t / a^2} - 1)}{\xi_n^2 (\lambda - \xi_n^2 D / a^2)} \quad (6)$$

The values  $\xi_n$  are the roots of the equation  $J_0(x) = 0$ , where  $J_0(x)$  is the Bessel function of the first kind of zero order. The analogous solution for the infinite plate is

$$\frac{\bar{C}}{N} = \frac{8}{\pi^2} \sum_{n=0}^{\infty} \frac{\lambda(e^{\lambda t - (2n+1)^2 \pi^2 D t / a^2} - 1)}{(2n+1)^2 [\lambda - (2n+1)^2 \pi^2 D / a^2]} \quad (7)$$

where  $a$  is the thickness of the plate.

Some ages resulting from diffusion losses from an infinite cylinder and an infinite plate are

given in Table 3. The calculations are for minerals having an age of 2800 m.y. If the results given in Table 3 are plotted on a diagram like Figure 1, the points define  $D/a^2$  curves which are indistinguishable from the one given in Figure 1 on the scale of the drawing. The only difference is that points corresponding to a particular value of  $D/a^2$  plot at different positions along the curve for different shapes. For a given thickness or diameter and value of  $D$ , a sphere will lose more lead than the infinite cylinder, which in turn will lose more lead than the infinite plate.

These calculations show that diffusion losses occurring in one, two, or three dimensions will produce essentially the same  $D/a^2$  values given in Figure 1. This indicates that the relationship depicted there is not likely to depend in any critical manner on the shape of the crystal units from which diffusion losses have occurred. For example, a zircon crystal may be approximated by a section of cylinder cut by two hemispheres. A rigorous solution to diffusion losses from such a body is obtained by



TABLE 3. Ages Resulting from Loss of Lead by Diffusion from Nonspherical Crystals of Age 2800 Million Years

| $a^{2*} \times 10^{-12} \text{ yr}^{-1}$ | Infinite Cylinder                     |                                       | Infinite Plate                        |                                       |
|--|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
|  | Pb <sup>206</sup><br>U <sup>238</sup> | Pb <sup>207</sup><br>U <sup>235</sup> | Pb <sup>206</sup><br>U <sup>238</sup> | Pb <sup>207</sup><br>U <sup>235</sup> |
|  | m.y.                                  | m.y.                                  | m.y.                                  | m.y.                                  |
| 6.25                                     | 2300                                  | 2545                                  | 2310                                  | 2535                                  |
| 12.50                                    | 2135                                  | 2435                                  | 2095                                  | 2400                                  |
| 25.0                                     | 1885                                  | 2260                                  | 1780                                  | 2175                                  |
| 50.0                                     | 1525                                  | 1985                                  | 1340                                  | 1785                                  |
| 100.0                                    | 1090                                  | 1550                                  | 835                                   | 1295                                  |
| 200.0                                    | 635                                   | 950                                   | ...                                   | ...                                   |

\* Note that  $a$  is the radius of the cylinder and  $e$  the total thickness of the plate.

combination of the infinite cylinder and the sphere calculations.

As was mentioned earlier, it is by no means certain that the physical size and shape of a crystal are the determining geometric factor for diffusion loss. In fact, some evidence exists to the contrary. Nicolaysen, de Villiers, Burger, and Kretz [1958] have reported uranium-lead ages for three different size fractions of zircon from Ottensville granite where mineral age studies indicate an age of about 2000 m.y. The <sup>238</sup>Pb-<sup>206</sup>Pb ages varied with grain size as follows:

100 mesh (average mesh opening 0.023 cm): 816 m.y.  
 60-170 mesh (average mesh opening 0.010 cm): 689 m.y.  
 0-325 mesh (average mesh opening 0.0066 cm): 642 m.y.

According to the authors, there is no evidence of any recent geologic process that might have caused the loss of lead. These results do not preclude the possibility that the effective radius for diffusion is of the order of the grain size, but they strongly suggest that the effective radius is not the physical radius.

**Activation energy for diffusion.** In several cases there is evidence to indicate that zircon has not lost appreciable lead during conditions of metamorphism. A good example occurs in the Baltimore gneiss at Baltimore, Maryland. Zircons from two samples of gneiss have isotopic lead ages which indicate that the minerals have

lost only 10 to 20 per cent of their lead over the past 1150 m.y., the presumed age of the minerals [Tilton, Wetherill, Davis, and Hopson, 1958]. Such a small loss of lead is found although the gneiss occurs in the Piedmont, where metamorphism was presumably intense as a result of processes associated with the formation of the Appalachians. The biotite age of 300 m.y. does in fact appear to be related to such processes.

The five Finnish zircons in Table 1 have discordant age patterns that are better described by diffusion losses that took place over the entire time the minerals have existed, even though biotite samples taken from the same rocks as the zircons give ages of 1800 m.y. (Wetherill, Kouvo, Tilton, and Gast, paper to be submitted to the *Journal of Geology*). These observations indicate that the diffusion rates may be relatively insensitive to temperature; that is, the activation energy for diffusion is low. If it is assumed that the Finnish zircons were subjected to temperatures of 400° to 600°C for 10 to 100 m.y. 1800 m.y. ago and have been at a temperature of approximately 50°C for most of the remainder of the time, calculations using the relation

$$D(T) = D_0 \exp(-Q/RT)$$

permit an estimation of the activation energy for diffusion,  $Q$ . The samples from Heinävaara and Sotkuma show the most evidence of possible loss of lead at 1800 m.y., but even these samples appear to have lost no more than 30 per cent of their lead at that time. Such a loss at 400° to 600°C coupled with additional required losses at 50°C give values of 3 to 10 kcal/mol for the activation energy (in the temperature range 50° to 600°C). The fact that zircons from shield areas like southern Ontario appear to have lost about 10 per cent of their lead by diffusion again argues for a low value of the activation energy, since there is no evidence for metamorphism in this area in the last 1000 m.y.

Activation energies observed for volume diffusion in metals vary from 20 to 80 kcal/mol and average about 30 kcal/mol [Jost, 1952, pp. 234-237]. Limited data for diffusion in ionic crystals indicate that activation energies are somewhat lower than for metals [Jost, 1952, p. 199; Barrer, 1951, p. 274]. Values of 35 to 85 kcal/mol are quoted in the literature for the diffusion of argon from micas [Gerling and



Morozova, 1957; Amirkhanov, Brandt, Bartnitskii, Gasanov, and Gurvich, 1959; Fechtig, Gentner, and Zähringer, 1960]. By comparison, the estimated activation energy for lead diffusion is surprisingly low, but qualitative considerations indicate the possibility of the lead value.

Diffusion in a lattice is generally considered to depend on the presence of vacancies. Then the term  $Q$  in the equation

$$D(T) = D_0 \exp(-Q/RT)$$

can be composed of two parts,  $E$  and  $U$  [Jost, 1952, p. 135].  $E$  is the energy required to form 1 mol of vacancies in the lattice, and  $U$  is the energy barrier to be surmounted by a diffusing particle in moving from the vicinity of a vacancy. Since crystals are imperfect, they contain some vacancies as a result of impurities and imperfections. These are temperature-independent vacancies. Thus a plot of  $\log D$  against  $1/T$  may show a break in slope (knee) at a temperature that depends on the number of temperature-independent vacancies present. Ideally (for a simple, one-component lattice), the high-temperature branch is described by

$$D(T) = D_0 \exp(-[U + E]/RT)$$

and the lower branch by

$$D'(T) = D'_0 \exp(-U/RT)$$

Examples fitting this hypothesis are the diffusion of sodium in NaCl and NaBr [Mapother, Crooks, and Maurer, 1950]. Fechtig, Gentner, and Zähringer [1960] show examples for the diffusion of argon in feldspar, anorthite, augite, and margarite, where knees in the curves occur at temperatures of 150° to 350°C. Amirkhanov, Brandt, Bartnitskii, Gasanov, and Gurvich [1959] find an increase in the temperature dependence of  $D$  for argon diffusion from a muscovite at 600°C.

The effect of temperature-independent vacancies in crystals has also been illustrated from conductivity data. Since summaries of these data are given by Jost [1952, chapter 4] and by Barrer [1951, chapter 6], only a brief discussion will be presented here. Of particular interest are the experiments of Koch and Wagner [1937], who have shown that the addition of 0.1 per cent of CdCl<sub>2</sub> to AgCl decreases the temperature dependence of conductivity compared with that

of pure AgCl below a temperature of 300°C. They also show that, for a plot of  $\log \sigma$  against  $1/T$ , the curve for AgBr plus 0.5 per cent PbBr<sub>2</sub> lies above that for AgBr plus 0.3 per cent PbBr<sub>2</sub> in the low-temperature region. The change to lower dependence on temperature occurs at approximately 300°C for 0.5 per cent PbBr<sub>2</sub> and 280°C for 0.3 per cent PbBr<sub>2</sub>. This is expected since 0.5 per cent PbBr<sub>2</sub> would create more vacancies in the lattice than 0.3 per cent PbBr<sub>2</sub> if electrical neutrality is to be maintained. At low temperatures the energy barrier for ion migration is found to be 25 kcal/mol for AgCl and 15 kcal/mol for AgBr. The corresponding values at low temperatures are 6.5 kcal/mol for AgBr plus CdCl<sub>2</sub> and 8.2 kcal/mol for AgBr plus PbBr<sub>2</sub>. These experiments clearly demonstrate the importance of temperature-independent vacancies on the mobility of ions in a crystal lattice.

The activation energy for lead diffusion in zircon was estimated over a temperature range of 50° to 600°C. It is possible that temperature-independent vacancies control diffusion in this range, hence the low value of activation energy. Radiation damage would be expected to contribute to the supply of temperature-independent vacancies. However, if it controlled the supply, the diffusion coefficient would increase as the amount of bombardment increased. It seems unlikely that the relationships shown in Figures 1 to 4 would result from such a process.

Solubility is still another factor that influences the value of the activation energy for diffusion in lead, it is observed that silver in gold, which have low solubilities, diffuse with lower activation energies (and with faster rates) at low temperatures than metals such as copper and tin, which have greater solubilities. These relationships are illustrated by Barrer [1951, p. 288]. A related fact is that activation energies for diffusion in metals are higher for self-diffusion than for foreign metals in alloys. Examples given by Jost [1952, pp. 234-235]. These data indicate that species with low solubilities in a lattice diffuse more readily than those with high stabilities. Lead has a low solubility in zircon, this being one of the principal reasons why zircon has been chosen over other minerals for determining ages by the diffusion methods.

TABLE 4. Comparison of Zircon and Biotite Ages\* from Two Rocks

| Rock               | Mineral | Age, million years                       |  |   |   |                               |                             |
|--------------------|---------|--|--|---|---|-------------------------------|-----------------------------|
|                    |         | $\frac{\text{Pb}^{206}}{\text{U}^{238}}$ | $\frac{\text{Pb}^{207}}{\text{U}^{235}}$ | $\frac{\text{Pb}^{207}}{\text{Pb}^{206}}$ | $\frac{\text{Pb}^{208}}{\text{Th}^{232}}$ | $\frac{\text{Rb}}{\text{Sr}}$ | $\frac{\text{K}}{\text{A}}$ |
| Baltimore gneiss   | Zircon  | 1040                                     | 1070                                     | 1120                                      | 940                                       |                               |                             |
|                    | Biotite |  |  |   |   | 305                           | 340                         |
| Pikes Peak granite | Zircon  | 624                                      | 707                                      | 980                                       | 313                                       |                               |                             |
|                    | Biotite |  |  |   |   | 1020                          | 980                         |

$$* \lambda(\text{Rb}^{87}) = 1.39 \times 10^{-11} \text{ yr}^{-1}$$

$$\lambda_e(\text{K}^{40}) = 0.583 \times 10^{-10} \text{ yr}^{-1}$$

$$\lambda_g(\text{K}^{40}) = 4.72 \times 10^{-10} \text{ yr}^{-1}$$

It is possible that the lead ages could be reconciled with an activation energy of zero. In this case recoil and/or collision phenomena might be considered as the cause for the loss of lead. *Russell and Ahrens* [1957] proposed several mechanisms whereby the loss of greater proportions of  $\text{Pb}^{207}$  than  $\text{Pb}^{206}$  might be explained in terms of differences in recoil energies of various intermediate daughter products in the two uranium decay series. It is difficult to understand how such processes could produce losses of radiogenic lead of 10 to 50 per cent as required for most of the minerals in Table 1. Alpha decay imparts energies of the order of 100,000 electron volts to the parent recoil atom. The range-energy relationship for fission fragments reported by *Bøggild* [1941] shows that such an atom would have a range of  $\sim 4 \times 10^{-6}$  cm in a mineral. After emission of 8  $\alpha$  particles, the average  $\text{U}^{238}$  nucleus would have migrated about  $3.2 \times 10^{-5}$  cm. If zircon crystals are considered to be spheres of radius  $4 \times 10^{-8}$  cm, corresponding to a size that would pass a 200-mesh sieve, it is found that only a few tenths of a per cent of lead can escape from the crystals by recoil.

If recoil energy is used only to transport the lead atoms to grain boundaries in order to permit escape of a greater proportion, the loss will be controlled by grain boundary diffusion. This in itself should be capable of explaining the discordant age patterns, since diffusion along grain boundaries would be essentially two-dimensional and ought to be somewhat analogous to the case of the infinite cylinder discussed earlier. Such a process would then give the same shape of 'discordia' curves derived for volume diffusion from the sphere.

Another mechanism having zero activation energy is to assume that the probability that a lead atom will be ejected from a crystal is proportional to the total bombardment in the crystal after the atom is formed. It does not seem possible to remove any substantial amount of lead by this means, and calculations show that, if such a process did control lead loss, the loss of  $\text{Pb}^{207}$  would be weighted too heavily to fit the discordant age patterns shown in Figures 1 to 4. The discordia curves appear to have unique characteristics that cannot be reproduced by simple models involving solely collision or recoil phenomena.

Diffusion may explain a difficulty concerning the behavior of the ages given by coexisting mica and zircon. The problem is illustrated by two examples in Table 4, taken from the work of *Tilton, Davis, Wetherill, and Aldrich* [1957] and *Tilton, Wetherill, Davis, and Hopson* [1958]. At Pikes Peak zircon has lost lead under conditions that do not appear to have affected the argon or strontium contents of biotite appreciably. In the Baltimore gneiss, biotite either crystallized or lost strontium and argon under conditions that did not seriously affect the lead in the zircon. In unmetamorphosed rocks, the uranium-lead ages of zircon are equal to or less than the rubidium-strontium and potassium-argon ages of micas. Here diffusion rates might be controlled mainly by the number of temperature-independent vacancies in the two mineral lattices. The results for the Baltimore gneiss can be explained by assuming that the temperature dependence in the interval 50° to 600°C is greater for the diffusion of argon and strontium in the biotite than for lead in the zircon. The

curves for log  $D$  as a function of  $1/T$  given by Fechtig, Gentner, and Zähringer [1960] show breaks in slope at temperatures of  $150^{\circ}$  to  $350^{\circ}\text{C}$ . Above these temperatures  $D$  increases sharply with  $T$ , presumably because temperature-dependent vacancies begin to predominate over temperature-independent vacancies in the lattice.  $D$  has a value of  $\sim 10^{-15}$   $\text{cm}^2/\text{sec}$  at  $500^{\circ}\text{C}$  for the four minerals studied. In  $10^7$  years, argon atoms would migrate  $\sim 0.8$  cm at this temperature. If the diffusion of lead in zircon is controlled by temperature-independent vacancies up to some temperature above  $500^{\circ}$  to  $600^{\circ}\text{C}$ , it should be possible to remove argon from biotite without affecting the lead content of the zircon to any great extent. Since there are but a few cases in the literature for which the rubidium-strontium age differs very greatly from the potassium-argon age of a mica, it appears that strontium diffusion is somewhat similar to argon diffusion in micas.

#### CONCLUSIONS

It has been shown that volume diffusion offers a promising explanation to several problems that have arisen in geochronology. It explains regularities that exist in the discordant uranium-lead ages of minerals from several continents. It does away with the need for episodic loss of lead in several cases where no evidence exists for metamorphic events at the time required. Moreover, it appears that volume diffusion will play a prominent part in resolving problems that have arisen from comparisons of uranium-lead ages of zircon with rubidium-strontium and potassium-argon ages of coexisting micas in granites and gneisses.

A fortunate aspect of the diffusion hypothesis is that it is amenable to rather rigorous testing even in the absence of laboratory determinations of the diffusion rates. Once the age of a group of rocks can be established by concordant uranium-lead ages or by other methods, the pattern of the discordant lead ages that will result from diffusion is determined. The importance of diffusion can then be judged from the number of cases for which discordant ages will fit the theoretical patterns. It would be particularly convincing to show that suites of minerals of two different ages from the same area fit two sepa-

rate discordia curves for loss of lead by diffusion. It would be difficult to invoke an episodic mechanism in such a case.

*Acknowledgments.* I am grateful to my colleagues, S. P. Clark, Jr., G. L. Davis, T. C. Hollister, and G. W. Wetherill, for helpful discussions and criticisms concerning various phases of this work.

#### REFERENCES

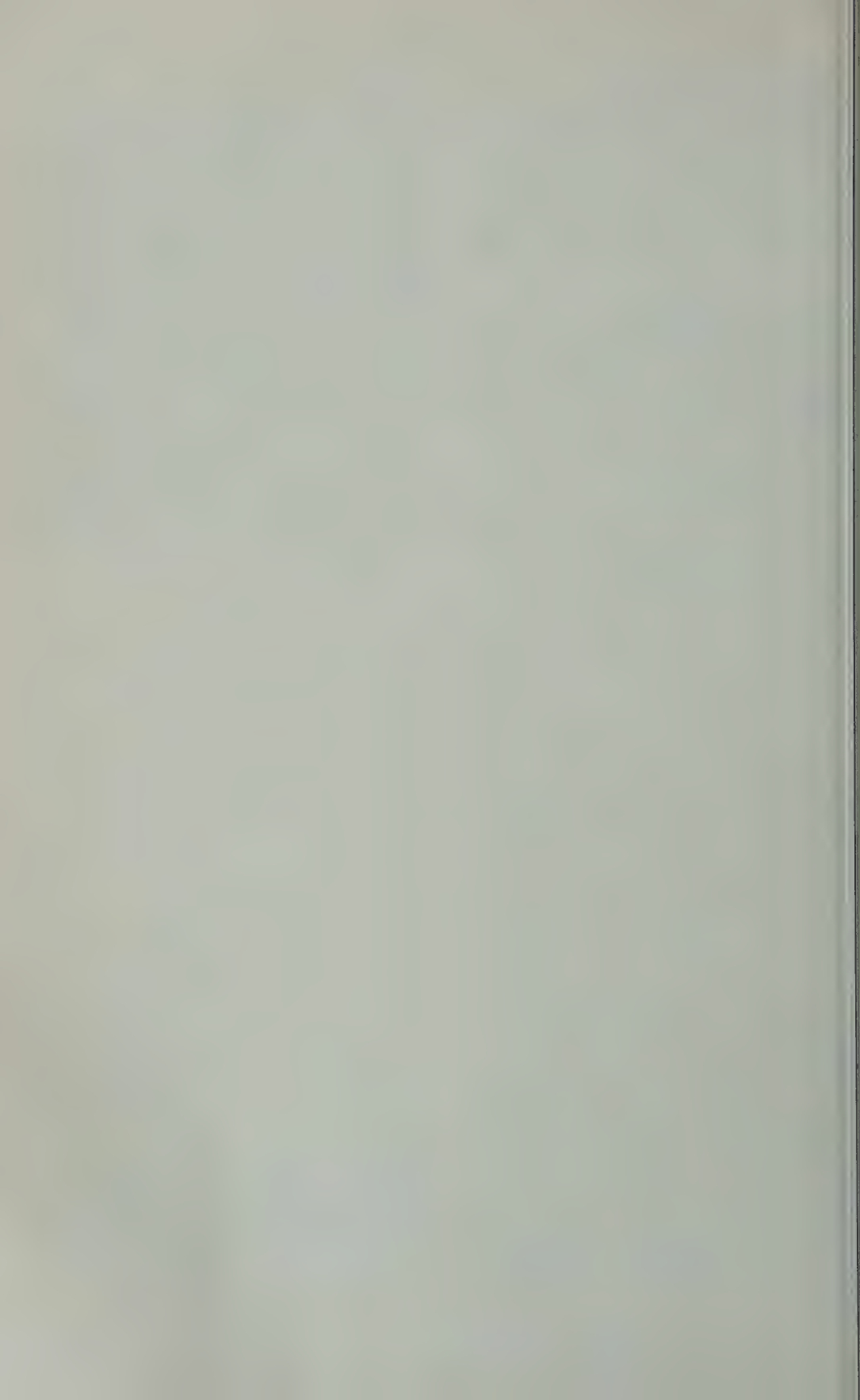
- Ahrens, L. H., The convergent lead ages of the oldest monazites and uraninites (Rhodesia, Manitoba, Madagascar, Transvaal), *Geochim. et Cosmochim. Acta*, **7**, 294-300, 1955a.  
 Ahrens, L. H., Implications of the Rhodesian pattern, *Geochim. et Cosmochim. Acta*, **8**, 1955b.  
 Aldrich, L. T., and G. W. Wetherill, Geochronology by radioactive decay, *Ann. Rev. Nucl. Sci.*, **8**, 257-298, 1958.  
 Amirkhanov, K. I., S. B. Brandt, E. I. Bartnik, S. A. Gasanov, and V. S. Gurvich, On the mechanism of losses of radiogenic argon in mica, *Izvest. Akad. Nauk SSSR, Ser. Geol.*, 1959.  
 Barrer, Richard M., *Diffusion in and through Solids*, University Press, Cambridge, 1951.  
 Bøggild, J. K., Range-velocity relation for fission fragments in helium, *Phys. Rev.*, **60**, 827-831, 1941.  
 Eckelmann, W. R., and J. L. Kulp, Uranium-lead method of age determination, 1, Lake Umbagog baskin problem, *Bull. Geol. Soc. Am.*, **67**, 335-356.  
 Fechtig, H., W. Gentner, and J. Zähringer, Disproportionsverluste von Argon in Mineralien und die Auswirkung auf die Kalium-argon Altersbestimmung, *Geochim. et Cosmochim. Acta*, **19**, 761-766, 1960.  
 Gast, P. W., J. L. Kulp, and L. E. Long, Absolute age of early Precambrian rocks in the Big Horn Basin of Wyoming, Montana and southeastern Manitoba, *Trans. Am. Geophys. Union*, **39**, 334, 1958.  
 Gerling, E. K., and I. M. Morozova, Determination of the activation energy of argon isolation from micas, *Geokhimiya*, 304-311, 1957.  
 Greenhalgh, D., and P. M. Jeffery, A contribution to the Precambrian chronology of Australia, *Geochim. et Cosmochim. Acta*, **16**, 39-57, 1952.  
 Holmes, A. H., The oldest dated minerals of the Rhodesian shield, *Nature*, **173**, 612, 1954.  
 Holmes, A. H., Dating the Precambrian of Peninsular India and Ceylon, *Proc. Geol. Assoc. Can.*, **7**, 81-106, 1955.  
 Holmes, A. H., and Lucien Cahen, *Géochronologie africaine*, *Mém. acad. roy. sci. coloniales*, 1957.  
 Jost, W., *Diffusion in Solids, Liquids, Gases*, Academic Press, New York, 1952.



- sch, E., and C. Wagner, Der Mechanismus der Ionenleitung in festen Salzen auf Grund von Fehlordnungsvorstellungen, I. Z., *physik. Chem.*, **3**, 38, 295-325, 1937.
- Suovo, O., Radioactive age of some Finnish Precambrian minerals, *Geol. Comm. Finland, Bull.*, **182**, 70 pp., 1958.
- Thompson, D., H. N. Crooks, and Robert Maurer, Self diffusion of sodium in sodium chloride and sodium bromide, *J. Chem. Phys.*, **18**, 1231-1236, 1950.
- Tröelsæien, L. O., Solid diffusion in radioactive minerals and the measurement of absolute age, *Geochim. et Cosmochim. Acta*, **11**, 41-59, 1957.
- Tröelsæien, L. O., J. W. L. de Villiers, A. J. Burger, and F. W. E. Strelow, New measurements relating to the absolute age of the Transvaal system and of the Bushveld igneous complex, *Trans. Geol. Soc. S. Africa*, **61**, 137-163, 1958.
- Urey, A. O., The isotopic constitution of radiogenic leads and the measurement of geological time, *Phys. Rev.*, **55**, 153-163, 1939.
- Wassell, R. D., and L. H. Ahrens, Additional regularities among discordant lead-uranium ages, *Geochim. et Cosmochim. Acta*, **11**, 213-218, 1957.
- Wetherill, I. E., The role of secondary processes in age determination by radiometric methods, *Geokhimiya*, 18-29, 1956.
- Tilton, G. R., G. L. Davis, G. W. Wetherill, and L. T. Aldrich, Isotopic ages of zircon from granites and pegmatites, *Trans. Am. Geophys. Union*, **38**, 360-371, 1957.
- Tilton, G. R., G. W. Wetherill, G. L. Davis, and C. A. Hopson, Ages of minerals from the Baltimore gneiss near Baltimore, Maryland, *Bull. Geol. Soc. Am.*, **69**, 1469-1474, 1958.
- Vinogradov, A. P., Comparison of data on the age of rocks obtained by different methods and geological conclusions, *Geokhimiya*, 3-17, 1956.
- Wasserburg, G. J., Argon<sup>40</sup>: potassium<sup>40</sup> dating, in *Nuclear Geology*, edited by H. Faul, John Wiley & Sons, New York, pp. 341-349, 1954.
- Wetherill, G. W., An interpretation of the Rhodesia and Witwatersrand age patterns, *Geochim. et Cosmochim. Acta*, **9**, 290-292, 1956a.
- Wetherill, G. W., Discordant uranium-lead ages, *Trans. Am. Geophys. Union*, **37**, 320-326, 1956b.
- Zhirov, K. K., Data summarized in E. K. Gerling, The influence of metamorphism on the results of age determination according to lead, *Geokhimiya*, 288-295, 1958.

(Manuscript received June 17, 1960.)





## The Distribution of Cosmic-Ray-Produced Rare Gases in Iron Meteorites

P. SIGNER AND A. O. NIER

*School of Physics, University of Minnesota  
Minneapolis 14, Minnesota*

**Abstract.** The cosmogenic  $\text{Ar}^{38}$ ,  $\text{Ar}^{39}$ ,  $\text{Ne}^{22}$ ,  $\text{Ne}^{21}$ ,  $\text{Ne}^{20}$ ,  $\text{He}^4$ , and  $\text{He}^3$  distribution in part of a cross section of the iron meteorite Grant has been measured. The amounts were found to vary systematically with depth. The depth effect was analyzed in terms of a production mechanism, and numerical parameters were determined. Comparisons were made with the results of other investigations. Functional relationships correlating relative amounts of cosmogenic nuclides with position and size of a meteoritic body are demonstrated. In principle the determination of a sufficient number of cosmogenic nuclides in a single sample permits deductions about the radiation dosage and the size of the meteoroid as well as the location of the sample within the body.

### INTRODUCTION

Paneth, Reasbeck, and Mayne [1952] were able to demonstrate that meteorites contained helium generated by cosmic-ray bombardment as predicted by Bauer [1947, 1948] and Huntley [1948]. According to current thinking the interaction of the energetic cosmic rays (principally protons having energies in the Bev range) with nuclei in meteorites causes spallation and fragmentation, with the result that virtually all possible nuclides lighter than the target are formed. The actual formation process may be quite complex inasmuch as the struck nucleus emits both high- and low-energy nucleons and mesons. These secondary particles, in turn, interact with other nuclei and thus contribute to the cosmogenic-produced nuclides observed.

A complete quantitative understanding of the production mechanism would enable one to determine from the amounts of cosmogenic nuclides observed information about a meteorite such as the radiation dosage received, the pre-atmospheric size of the body, and the position of a particular sample in the body relative to the surface. One can learn something further about the energy spectrum of primary cosmic rays and the constancy of intensity in space and time. The information and experimental data available today are not sufficient to formulate the complicated production mechanism precisely. Thus systematic experimental approaches can

be very helpful. Two possible lines of attack appear fruitful: a comparison of the amounts of cosmogenic nuclides found in a large number of samples chosen from different meteorites; or a complete investigation of depth variation in a single meteorite. This paper will deal with the latter approach.

A first attempt to verify the variation of the He content with depth was made on the meteorite Carbo by Paneth, Reasbeck, and Mayne [1953]. Fireman [1958] investigated the  $\text{He}^3$  and Hoffmann and Nier [1959] the  $\text{He}^3$  and  $\text{He}^4$  in the same meteorite. These same authors, Hoffmann and Nier [1958] and Fireman [1959], have also studied the helium in a cross section of the meteorite Grant. Keen Mountain and Casas Grandes were systematically analyzed for  $\text{He}^3$  and  $\text{He}^4$  by Hoffman and Nier [1960]. Vinogradov, Zadorozhnyi, and Florenskii [1957] in their study of the large iron meteorite Sikhotealin concluded that the Ar, Ne, and He contents exhibited a depth effect. On the other hand, Gerling and Levskii [1958], on the basis of their study of the same meteorite, questioned whether there actually is a variation of these cosmogenic nuclides with depth.

Fechtig, Genter, and Kistner [1960] attempted a study of the depth variation of Ar, Ne, and He in the iron meteorite Treysa and found some variations, but they were unable to make a detailed analysis.

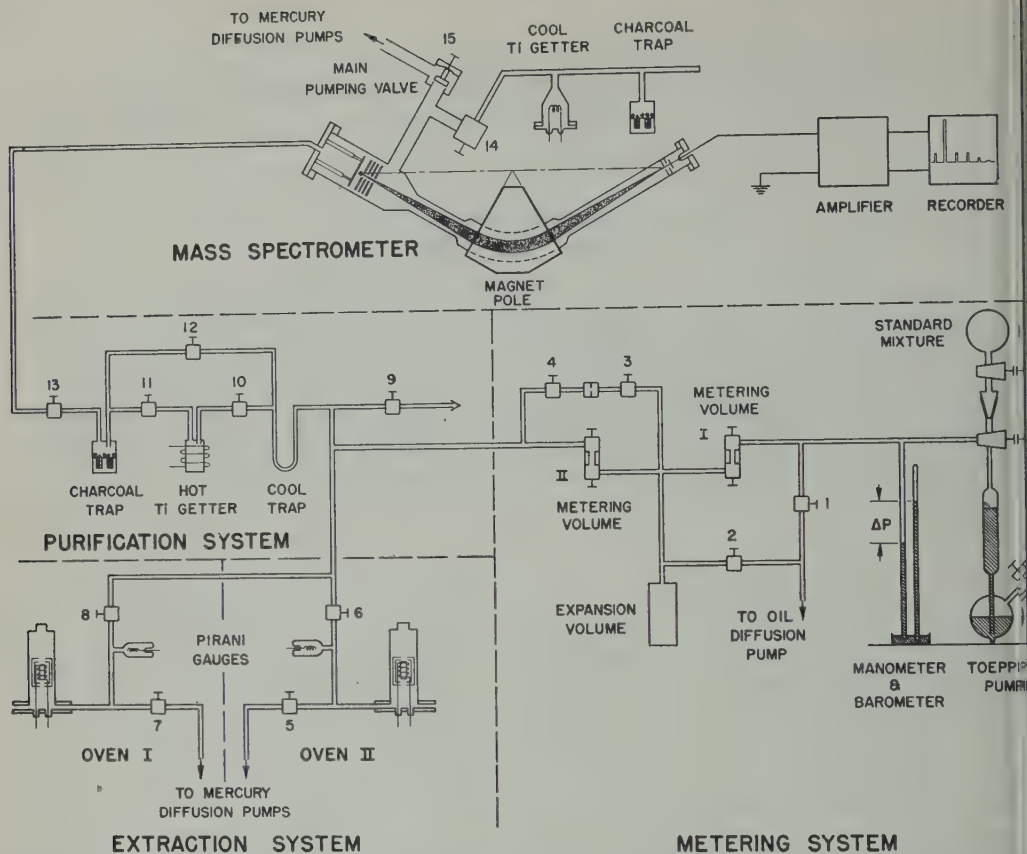


Fig. 1. Mass spectrometer and associated extraction, purification, and metering systems. Instruments can be used dynamically (for example, to test purity of standard mixture) by closing metering volume II and valves 6, 8, 10, and 11, and opening valves 3, 4, 12, and 13, permitting gas to flow through pinhole directly into mass spectrometer tube. Valve 9 permits samples extracted in other apparatus to be attached. Normally valve 9 is kept open and line is capped as shown. As schematically indicated, the ion currents are measured directly; i.e., no electron multiplier is used.

The present paper reports on an investigation of the  $\text{Ar}^{38}$ ,  $\text{Ar}^{39}$ ,  $\text{Ne}^{20}$ ,  $\text{Ne}^{21}$ ,  $\text{Ne}^{30}$ ,  $\text{He}^4$ , and  $\text{He}^3$  in the iron meteorite Grant. This meteorite was chosen because of the extremely favorable pattern of  $\text{He}^3$  and  $\text{He}^4$  distribution shown in the earlier investigations. The experimental results are analyzed in terms of the model for the production mechanism proposed by *Ebert and Wänke* [1957] and based on the ideas presented by *Martin* [1953].

### 1. APPARATUS AND PROCEDURE

Figure 1 is a schematic drawing showing the mass spectrometer and associated extraction, purification, and metering systems. The mass

spectrometer is a conventional  $60^\circ$  Nier type having a radius of curvature of 6 inches (15 cm). The spectrometer tube is constructed entirely of stainless steel parts which are either welded or silver-soldered together. The source and collector are assembled on flanges bolted to the tube, copper gaskets being used to make the vacuum seals. An all-metal valve (15) permits the spectrometer tube to be isolated from the pumping system so that it can be run statically, thus giving a very high sensitivity. An appendage which may be isolated by means of a second, similar valve (14) has a volume which has in it a tungsten filament wrapped with titanium wire. This may be flashed from time to time

cover the inner wall of the volume with a fresh layer of titanium that acts as a 'getter' for impurity gases. The appendage also has a charcoal trap which, when cooled with liquid nitrogen, permits removal of the argon from the neon and helium. Because of the construction, the entire system, mass spectrometer plus appendage, can be baked at 300°C.

The lower part of the figure shows the extraction, purification, and metering systems. All to the left of the metering volume I, used for admitting known amounts of standard mixtures during the calibration procedure, is also of all-metal construction, and so this, too, can be thoroughly baked at temperatures up to 300°C to remove impurities. The two independent gas extraction ovens, shown in the figure, permit preliminary outgassing of one set of samples while measurements are in progress on those extracted in the other oven.

The direct connection between the extraction system and the mass spectrometer tube is of great importance in working with small samples. First of all, the mass spectrometer can be used to check the leak tightness of the entire system immediately before the gas is released from the sample. Second, contamination or loss of sample—such as would be encountered if the sample were first collected in a bulb and later attached to the mass spectrometer for analysis—is completely avoided.

In the present investigation samples of 200 to 400 mg were sawed from the same bars used in the earlier helium work. The samples were washed in dilute HNO<sub>3</sub> and in a detergent solution before they were weighed and placed in the extraction system. The crucible for the extraction is made of recrystallized alumina and has a volume of 1 cm<sup>3</sup>. It is surrounded by a closely wound, closely spaced spiral made of 0.040-inch-diameter tungsten wire. This is placed in a second crucible, which in turn is surrounded by three concentric radiation shields. A resistance-heated furnace has an important advantage over the commonly used induction heating system: in the latter, if the pressure rises above a critical value during the heating process, an electrical discharge may take place, with the result that part of the gas will be lost through ion bombardment of surfaces.

Samples are not placed directly in the cruci-

ble. Instead, as many as six may be stored in an assembly above the crucible and dropped as needed by means of a magnetic release. This saves considerable time, inasmuch as a very tedious outgassing procedure with extensive heating of the entire extraction system must be followed each time the system is opened to air. Without this rigorous procedure the atmospheric argon contamination cannot be reduced to the very low level desirable in these experiments. During the outgassing process the samples are kept cool.

When an extraction is to be made, the extraction and purification systems are isolated from the pumps, the metering system and the mass spectrometer. For example, if oven I is used, valves 4, 6, 7, 12, 13, and the lower valve of metering volume II are closed; 8, 9, 10, 11, and the upper valve of metering volume II are open. The sample is dropped into the hot crucible, and it melts within 30 to 60 seconds. Complete evaporation takes place in less than 20 minutes. During this process those impurities not removed by the dry-ice-cooled cold trap are taken up by the hot titanium getter. Twenty minutes after the sample is dropped the crucible and titanium-getter heaters are shut off, and cooling takes place for 10 minutes.

Valve 15 is now closed, 13 is opened; and the sample expands into the mass spectrometer tube. Measurements are first made on the argon isotopes. The charcoal traps are then cooled to liquid-nitrogen temperature, which removes essentially all the argon, a desirable step, since doubly charged Ar<sup>40</sup> interferes with the measurement of Ne<sup>20</sup>. Neon and helium are then measured, and the whole system is evacuated. The entire measurement procedure requires approximately 1 hour.

Mass positions corresponding to possible disturbing impurities are observed in addition to those positions corresponding to the rare gas isotopes. In practice, only for Ne<sup>22</sup> and Ne<sup>30</sup> are corrections for impurities sometimes required (doubly charged CO<sub>2</sub> and Ar<sup>40</sup>, respectively). Such corrections did not exceed 1 or 2 per cent in the present investigation.

To determine the absolute amounts of the various rare gas isotopes observed, the mass spectrometer is calibrated both immediately before and after a sample analysis by the same



procedure as in the sample run. The calibration is performed with a standard mixture of  $\text{He}^3$ ,  $\text{He}^4$ , and normal neon and argon. The proportions of the four components are chosen to correspond very closely to the relative amounts of  $\text{He}^3$ ,  $\text{He}^4$ , neon, and argon observed in the samples under investigation. Known amounts of the standard mixture, approximately the same as the sample under investigation, can be admitted to the system through use of the manometer and the accurately known metering volumes I and II. The known isotopic composition of the standard mixture serves to determine the mass discrimination of the mass spectrometer; calibration before and after a run provides a check of the long-time stability of the mass spectrometer.

After the completion of the second calibration, the system is again evacuated. The sample crucible is then heated, the same cycle of events being followed as in the initial extraction. This serves to test the completeness of the extraction and also provides an additional blank run. The argon yield is generally less than 0.5 per cent of the earlier value, and the neon and helium are completely negligible.

Although not explicitly stated above, it should be understood that before the start of the actual calibration and sample analysis a blank run was made. This included all steps of a regular sample run except that no sample was dropped into the hot crucible. Only when the blank showed that the system was tight and clean did the run proceed.

## 2. EXPERIMENTAL RESULTS

Grant is a 480-kg fine octahedrite iron meteorite found in New Mexico. It is in the possession of the U. S. National Museum in Washington, D. C. A slab approximately 1 cm thick, passing through the center of mass, was removed and cut into bars. Through the courtesy of E. P. Henderson the bars designated as B, F, J, N, and R were placed at our disposal.  $\text{He}^3$  and  $\text{He}^4$  analyses on these have already been reported by Hoffman and Nier [1958].

In the present work samples were chosen to include the regions which, according to the earlier work, had the lowest concentration and the biggest gradient in the concentration of helium. Of

the 35 measurements made, those for bar F are shown in Table 1. The ratios  $\text{Ar}^{36}/\text{Ar}^{38}$ ,  $\text{Ne}^{20}/\text{Ne}^{22}$ , and  $\text{He}^3/\text{He}^4$ , corrected for mass discrimination, are found by averaging the measurements on six spectra. The standard deviation in these averages is generally less than 0.3 per cent. The absolute amounts of  $\text{Ar}^{38}$ ,  $\text{Ne}^{21}$ , and  $\text{He}^3$  are found by extrapolating the peak heights for these isotopes back to the time of admission of the sample and comparing values found with the corresponding values for the standard mixture in the calibration runs.

In most cases the samples sawed from the bars were cut in half, the halves being called *a* and *b* respectively. In three cases—samples 15, and 56—both halves were run. The comparison of the results in these three cases gives a measure of the reproducibility of the entire procedure. The reliability of the data may also be seen from the plots in Figure 5. In comparing two samples we believe the relative amounts of a particular isotope are accurate to 5 per cent, whereas the absolute amounts may have an error of 7 per cent. Relative isotopic ratios are believed to be accurate to 2 per cent, and absolute ratios to 4 per cent.

In making analyses of small gas samples by the static method, one encounters experimental difficulties which are generally less important in dynamic measurements. Through ion pumping and absorption on the clean surfaces—predicted ion pumping in our case—gas is lost. Conversely, gas may be released from the surfaces of the system when a new sample is introduced. In our procedure the ion pumping effect prevails, with the result that the peak heights are extrapolated back to the time of admission of the sample to the mass spectrometer. The rate of variation with time of argon is considerably higher than that of neon or helium.

Isotopic ratios in neon and helium do not vary with time, and hence need not be extrapolated back to zero time. On the other hand, the  $\text{Ar}^{38}$  varies slightly with time as a result of the slow release of absorbed atmospheric argon from the surfaces. When correction for this effect is applied, the  $\text{Ar}^{36}/\text{Ar}^{38}$  ratio remains constant with time.

(a) *Argon.* Figure 2 shows an argon spectrum taken 14 minutes after sample 56b was admitted to the mass spectrometer. The absolute

TABLE 1. Experimental Results for Samples Taken from Bar F of Grant Meteorite

| 1             | 2          | 3            | 4  | 5  | 6  | 7  | 8  | 9  | 10                                      | 11                                      | 12                        |
|---------------|------------|--------------|--|--|--|--|--|--|---|---|---------------------------|
| Sample Number | Weight, mg | Location, mm | $(\text{Ar}^{40}/\text{Ar}^{38})_0$<br>Extrapolated<br>to Time of<br>Admission | $\text{Ar}^{38} \times 10^8$ ,<br>cc STP/g | $\text{Ar}^{36}/\text{Ar}^{38}$<br>Corrected<br>for Air<br>Contamination | $\text{Ne}^{21} \times 10^8$ ,<br>cc STP/g | $\text{Ne}^{20}/\text{Ne}^{21}$<br>Not Corrected<br>for Air<br>Contamination | $\text{Ne}^{22}/\text{Ne}^{21}$<br>Not Corrected<br>for Air<br>Contamination | $\text{He}^4 \times 10^8$ ,<br>cc STP/g | $\text{He}^3 \times 10^3$ ,<br>cc STP/g | $\text{He}^3/\text{He}^4$ |
| S 32a         | 207        | -340         | 0.92   | 30.0                                       | 0.625  | 5.95                                       | 0.99   | 1.05   | 1880                                    | 490                                     | 0.261                     |
| S 14b         | 220        | -311         | 0.52   | 29.0                                       | 0.620  | 5.65                                       | 1.02   | 1.05   | 1840                                    | 465                                     | 0.253                     |
| S 14a         | 235        | -311         | 0.72   | 30.0                                       | 0.620  | 5.35                                       | ...  | ...  | 1900                                    | 485                                     | 0.255                     |
| S 29b         | 207        | -240         | 1.44   | 28.0                                       | 0.625  | 5.15                                       | 1.00   | 1.05   | 1860                                    | 455                                     | 0.245                     |
| S 31a         | 283        | -180         | 0.64   | 27.5                                       | 0.620  | 5.15                                       | 0.98   | 1.05   | 1780                                    | 450                                     | 0.253                     |
| S 56a         | 275        | -153         | 1.25   | 27.0                                       | 0.620  | 5.25                                       | 1.03   | 1.05   | 1785                                    | 450                                     | 0.252                     |
| S 56b         | 442        | -153         | 0.52   | 26.5                                       | 0.620  | 5.05                                       | 0.99   | 1.05   | 1770                                    | 435                                     | 0.245                     |
| S 54          | 376        | -149         | 1.66   | 27.5                                       | 0.620  | 5.30                                       | ...  | ...  | 1770                                    | 445                                     | 0.251                     |
| S 30a         | 373        | -112         | 0.64   | 27.0                                       | 0.620  | 5.15                                       | 0.98   | 1.05   | 1770                                    | 435                                     | 0.246                     |
| S 28b         | 201        | -60          | 0.98   | 27.5                                       | 0.625  | 5.10                                       | 1.00   | 1.05   | 1750                                    | 445                                     | 0.254                     |
| S 13b         | 289        | -6           | 0.91   | 27.5                                       | 0.615  | 5.20                                       | ...  | ...  | 1820                                    | 465                                     | 0.255                     |
| S 15b         | 240        | +63          | 0.57   | 28.5                                       | 0.625  | 5.85                                       | 1.02   | 1.05   | 1790                                    | 470                                     | 0.262                     |
| S 15a         | 262        | +63          | 1.25   | 30.0                                       | 0.620  | 5.70                                       | ...  | ...  | 1900                                    | 500                                     | 0.263                     |
| S 16a         | 290        | +132         | 0.56   | 31.0                                       | 0.635  | 5.90                                       | ...  | ...  | 1890                                    | 520                                     | 0.275                     |
| S 33a         | 294        | +158         | 1.14   | 31.0                                       | 0.640  | 7.05                                       | 1.01   | 1.04   | 1890                                    | 520                                     | 0.275                     |
| S 52a         | 157        | +159         | 2.40   | 31.0                                       | 0.645  | 7.05                                       | 1.06   | 1.05   | 2000                                    | 525                                     | 0.263                     |

\* Location (+) or (-) refers to distance, right or left respectively, from reference line scratched on cross-section slab of meteorite.

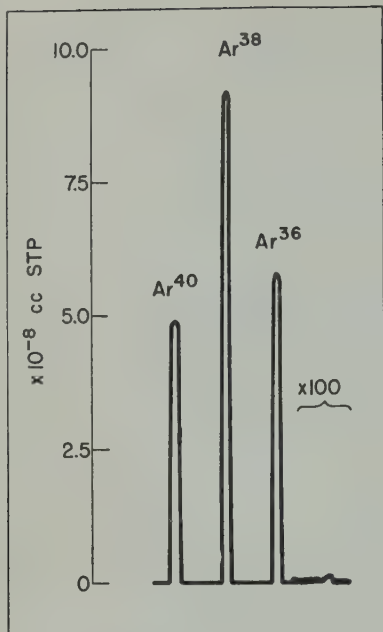


Fig. 2. Argon spectrum of sample 56b taken 14 minutes after sample was admitted to mass spectrometer.  $\text{Ar}^{38}$  peak is equivalent to  $3.5 \times 10^{-11}$  ampere. Impurity peaks in the mass positions 41, 39, and 37 have same magnitude as 34 impurity shown. Note that mass position 35 has no detectable residual.

of peaks in the odd mass positions indicates that corrections for impurities such as hydrocarbons or HCl are unnecessary. Column 4 of Table 1 gives the  $\text{Ar}^{40}/\text{Ar}^{38}$  ratios extrapolated back to time of admission of the sample to the mass spectrometer. Although these ratios are considerably lower than those reported in some investigations, we believe that they are still high because of the presence of some contaminating atmospheric argon. For example, in the analysis of sample 56b the blank run and the second extraction run indicate a contamination which, if allowed for, reduces the value for  $\text{Ar}^{40}/\text{Ar}^{38}$  from 0.52 to 0.4. Although it is tempting to attribute all the variation in column 4 to contamination effects, the conditions in some of the runs showing the relatively high  $\text{Ar}^{40}$  were so favorable that part of the variation may be due to variable radiogenic argon as suggested by *Stoenner and Zähringer* [1958]. If the variation were associated with variable amounts of primordial gas, one might expect a related variation in the

$\text{Ar}^{36}/\text{Ar}^{38}$  ratio. Such a variation could not be detected. Moreover, a search for krypton and xenon in Grant gave upper limits of abundance of less than  $10^{-10}$  cc STP/g. Because some of  $\text{Ar}^{40}$  is radiogenic, it appears that the  $(\text{Ar}^{40}/\text{Ar}^{38})_{\text{cosmogenic}}$  is less than 0.4. This is consistent with target studies now under way in this laboratory and with the work of *Schaepfer and Zähringer* [1959] and *Bieri* [1958].

Column 5 gives the  $\text{Ar}^{38}$  amounts actually observed. No correction for atmospheric argon contamination was necessary, even assuming that the  $\text{Ar}^{40}$  observed to be due to such contamination. The  $\text{Ar}^{36}/\text{Ar}^{38}$  ratios are computed with the assumption that all the observed  $\text{Ar}^{40}$  is due to atmospheric argon contamination. This is an overcorrection, since certainly some of the  $\text{Ar}^{40}$  is of meteoritic origin. Fortunately, the effect is very small, and in no case in Table 1 did the correction decrease the  $\text{Ar}^{36}/\text{Ar}^{38}$  ratio by more than 1 per cent.

(b) *Neon.*  $\text{Ne}^{21}$  is the best measure of cosmogenic neon because no corrections for contamination are necessary. The measurement of the other isotopes involves corrections such as doubly charged  $\text{CO}_2$  in the case of  $\text{Ne}^{22}$ . In the case of  $\text{Ne}^{20}$ , contaminations such as  $\text{H}_2\text{O}$ ,  $\text{H}_2$ , or doubly charged  $\text{Ar}^{40}$  may occur. In the present work the corrections could be made with high precision; in no case did the corrections affect the  $\text{Ne}^{20}/\text{Ne}^{21}$  and the  $\text{Ne}^{22}/\text{Ne}^{21}$  ratios by more than 2 per cent.

The measurement of cosmogenic neon is also complicated by the possibility of traces of atmospheric neon contamination. In the blank runs small amounts of atmospheric neon were sometimes observed. Furthermore, in some runs where abnormally large amounts of atmospheric argon were present (more than in any of the runs reported in Table 1) a relatively large amount of atmospheric neon was present also. Thus, in run S 44 (a sample from bar J)  $(\text{Ar}^{40}/\text{Ar}^{38})_0 = 20$  and the neon isotopes had relative abundances as follows:  $\text{Ne}^{20} : \text{Ne}^{21} : \text{Ne}^{22} = 1.55 : 1 : 1.09$ .

The  $\text{Ne}^{21}$  amounts are given in column 7, the  $\text{Ne}^{20}/\text{Ne}^{21}$  and  $\text{Ne}^{22}/\text{Ne}^{21}$  ratios in column 8 and 9, respectively. The variations in the  $\text{Ne}^{20}/\text{Ne}^{21}$  ratios are believed to be caused by variable amounts of atmospheric neon contamination. In view of the knowledge gained from blank runs



the second-extraction processes, it is believed that, for cosmogenic neon,  $\text{Ne}^{30}/\text{Ne}^{21} = 0.96 \pm 0.05$  and  $\text{Ne}^{23}/\text{Ne}^{21} = 1.05 \pm 0.04$ . *eri* [1958] gave the corresponding ratios in *arbo* to be 0.93 and 1.09, respectively. The average values obtained by *Wänke and Hinzenberger* [1958] in five iron meteorites were 0.94 and 1.02, respectively. O. A. Schaeffer and *Zähringer* (private communication) found the  $\text{Ne}^{30}$  and  $\text{Ne}^{23}$  amounts in three iron meteorites to be equal to or slightly larger than the  $\text{Ne}^{21}$ . For the iron meteorite *Treysa*, *Fechtig, Gentner, and Kistner* [1960] give a corrected  $\text{Ne}^{22}/\text{Ne}^{21}$  ratio of 1.08 but give no value for  $\text{Ne}^{30}/\text{Ne}^{21}$ , indicating merely that this ratio is larger than 1. *Snogradov, Zadorozhnyi, and Florenskii* [1957] reported that in *Sikhote-Alin*  $\text{Ne}^{30} : \text{Ne}^{21} : \text{Ne}^{23} = 1.14 : 1 : 1.04$ . *Gerling and Leviskii* [1958], working on the same meteorite, also reported  $\text{Ne}^{30}$  and  $\text{Ne}^{23}$  to be more abundant than  $\text{Ne}^{21}$ . In these last two publications the  $\text{Ar}^{40}$  was found to be 10 to 100 times as abundant as  $\text{Ar}^{38}$ .

For the iron meteorite *Thunda*, *Reasbeck and Mayne* [1955] in their pioneering work report  $\text{Ne}^{30} : \text{Ne}^{21} : \text{Ne}^{23} = 43 : 36 : 21$ . These results differ markedly from all the values just cited and have not been substantiated. On the other hand, in the iron meteorite *Washington County*, *Schaeffer and Fisher* [1959] found the  $\text{Ne}^{30}$  to be several times larger than the  $\text{Ne}^{21}$  or the  $\text{Ne}^{23}$ . The finding for this particular meteorite, which has been essentially confirmed by our own measurements, suggests the presence of primordial gas or gas from some source not associated with laboratory contamination.

(c) *Helium*. Columns 10 and 11 give the amounts of  $\text{He}^4$  and  $\text{He}^3$ , respectively. Blank runs showed that contributions to these amounts due to inleakage or outgassing were well below 1 per cent and hence negligible.  $\text{He}^3/\text{He}^4$  ratios are given in column 12. The present measurements are in close agreement with the corresponding ones found by *Hoffman and Nier* [1958].

TABLE 2. Experimental Results for Samples Taken from Bars J, N, and R of Grant Meteorite

| Sample Number | Location,* mm | $\text{Ar}^{36}/\text{Ar}^{38}$<br>Corrected<br>for Air<br>Contamination | $\text{Ar}^{38} \times 10^8$ ,<br>cc STP/g | $\text{Ne}^{21} \times 10^8$ ,<br>cc STP/g | $\text{He}^4 \times 10^8$ ,<br>cc STP/g | $\text{He}^3 \times 10^8$ ,<br>cc STP/g | $\text{He}^3/\text{He}^4$ |
|---------------|---------------|--|--|--|---|---|---------------------------|
| Bar J         |               |  |  |  |   |   |                           |
| 74            | -333          | 0.630  | 30.5                                       | 6.00                                       | 1890                                    | 490                                     | 0.259                     |
| 43            | -330          | 0.630  | 30.0                                       | 5.75                                       | 1850                                    | 485                                     | 0.262                     |
| 44            | -272          | 0.620  | 29.0                                       | 5.70                                       | ...                                     | ...                                     | ...                       |
| 49            | -194          | 0.625  | 28.5                                       | 5.40                                       | 1850                                    | 465                                     | 0.251                     |
| 50            | -137          | 0.620  | 29.0                                       | 5.40                                       | 1820                                    | 460                                     | 0.253                     |
| 46            | -104          | 0.625  | 27.5                                       | 5.40                                       | 1800                                    | 465                                     | 0.258                     |
| 51            | -61           | 0.625  | 28.5                                       | 5.50                                       | 1800                                    | 465                                     | 0.258                     |
| 42            | -58           | 0.620  | 28.0                                       | 5.45                                       | ...                                     | ...                                     | ...                       |
| 47            | -21           | 0.625  | 28.0                                       | 5.65                                       | 1830                                    | 475                                     | 0.260                     |
| 39            | +19           | 0.625  | 29.0                                       | 5.90                                       | ...                                     | ...                                     | ...                       |
| 40            | +75           | 0.635  | 30.5                                       | 6.25                                       | ...                                     | ...                                     | ...                       |
| 79            | +104          | 0.640  | 32.5                                       | 5.90                                       | 1880                                    | 520                                     | 0.277                     |
| 69            | +106          | 0.645  | 33.0                                       | 7.15                                       | ...                                     | ...                                     | ...                       |
| 68            | +113          | 0.645  | 33.0                                       | 7.30                                       | 1930                                    | 545                                     | 0.282                     |
| Bar N         |               |  |  |  |   |   |                           |
| 72b           | -95           | 0.625  | 28.5                                       | 6.00                                       | 1860                                    | 505                                     | 0.272                     |
| 90a           | +21           | 0.630  | 28.0                                       | 6.25                                       | 1810                                    | 500                                     | 0.276                     |
| 71a           | +64           | 0.640  | 34.5                                       | 7.45                                       | 1990                                    | 550                                     | 0.276                     |
| Bar R         |               |  |  |  |   |   |                           |
| 88            | -38           | 0.640  | ...  | 7.50                                       | 1930                                    | 555                                     | 0.288                     |
| 89            | +32           | 0.640  | ...  | 7.80                                       | 1950                                    | 560                                     | 0.287                     |

\* See note to Table 1.



## 3. DISCUSSION

(a) *Statistical analysis of the experimental results.* An examination of Table 1 shows that the amounts of the measured rare gases are lower at the center of the bar than at the ends. Since the experimental errors are not completely negligible compared with the variations caused by increasing shielding, it is clearly worth while first of all to analyze the measurements on a purely statistical basis. To make such an analysis objective we will neglect the geometric position of the sample in the body and look for correlations between amounts of two nuclides or between amounts of a particular nuclide and an isotopic abundance ratio.

Table 2 gives the experimental results for the bars J, N, and R. These results together with those given in Table 1 are the only data employed in making the analysis. It is to be noted that because of experimental difficulties in a few runs, data on some nuclides are missing.

Figure 3, *a*, *b*, and *c*, shows the respective amounts of  $\text{Ar}^{38}$ ,  $\text{He}^4$ , and  $\text{He}^3$  plotted against the corresponding amounts of  $\text{Ne}^{21}$ . If, for example, one assumes as a first approximation a linear relation between the amounts of  $\text{Ar}^{38}$  and  $\text{Ne}^{21}$ , one computes from a least-squares analysis that

$$\text{Ar}^{38} = (15.4 \pm 2.38\text{Ne}^{21}) \times 10^{-8} \text{ cc STP/g} \quad (1)$$

where the standard deviation for the  $\text{Ar}^{38}$  around this line is 3.6 per cent of the mean amount of  $\text{Ar}^{38}$ . On the other hand, if no correlation is assumed and merely an average value is computed for  $\text{Ar}^{38}$ , a value of  $29.3 \times 10^{-8}$  cc STP/g with a standard deviation of 6.6 per cent is found. This error is shown in the figure by the pair of dashed horizontal lines on either side of the heavier dashed line representing the average value. The difference in the errors clearly suggests a correlation in the relative amounts of the two nuclides in question.

Figure 3, *b* and *c*, indicates a similar effect when  $\text{He}^4$  and  $\text{He}^3$  are plotted against  $\text{Ne}^{21}$ , and Figure 4, *a* and *b*, shows the correlation of  $\text{Ar}^{38}/\text{Ar}^{39}$  with  $\text{Ar}^{38}$  and  $\text{He}^3/\text{He}^4$  with  $\text{He}^3$  respectively.

From the fact that the standard deviation  $\bar{S}$  (based on no correlation) is for all cases larger than  $\hat{S}$  (based on the assumption of a linear

correlation) it is concluded that the variables are systematically correlated. This in itself does not necessarily establish the existence of a depth effect. Only when a geometric analysis also yields a correlation can we conclude that a depth effect is present. Random loss of gases might result in a similar correlation.

The standard deviations in the amounts of  $\text{Ar}^{38}$ ,  $\text{He}^4$ , and  $\text{He}^3$  have approximately the same magnitude. This suggests that they are associated with experimental errors in the measurement rather than random fluctuation in the amounts of gas present in volumes of the order of the samples studied (cubes having sides approximately 4 mm). The case of  $\text{He}^4$  is particularly interesting since the presence of localized amounts of U and Th would lead to spots where the  $\text{He}^4$  concentration would be particularly high. Not much is known about the homogeneity of the U and Th distribution in meteorites; it is therefore interesting to note that a comparison of the standard deviations of  $\text{He}^3$  and  $\text{He}^4$  leads to the conclusion that the random variation in the  $\text{He}^4$  content is less than 1 per cent, or thereabouts. If it is assumed that the crystallization of the meteorite took place 4.5 aeons ago and that the Th/U ratio was unity, it would require an excess U content of  $10^{-10}$  in a small volume of the size considered here to give a 1 per cent excess amount of  $\text{He}^4$ .

(b) *Geometrical analysis of the experimental results.* To show that the correlations discussed in the previous section are really associated in a systematic way with the location of the sample in the meteorite, the amounts of various nuclides are plotted against position. Figure 5 shows a plot for bar F (data of Table 1). Similar graphs would be obtained for the other bars.

From data of this kind, contours of constant amounts can be drawn. Figure 6 shows contours for  $\text{Ne}^{21}$  and  $\text{Ar}^{38}$  for part of the section of the meteorite under investigation. In fitting the curves to the points, the assumption is made that the curves should have the same shape for the several nuclides. Within the experimental errors these curves agree with those given earlier for the helium isotopes by Hoff and Nier [1958] and by Fireman [1959].

From the topographic data of Figure 6 the radial increase was determined, the sloping line shown being employed as reference radius.

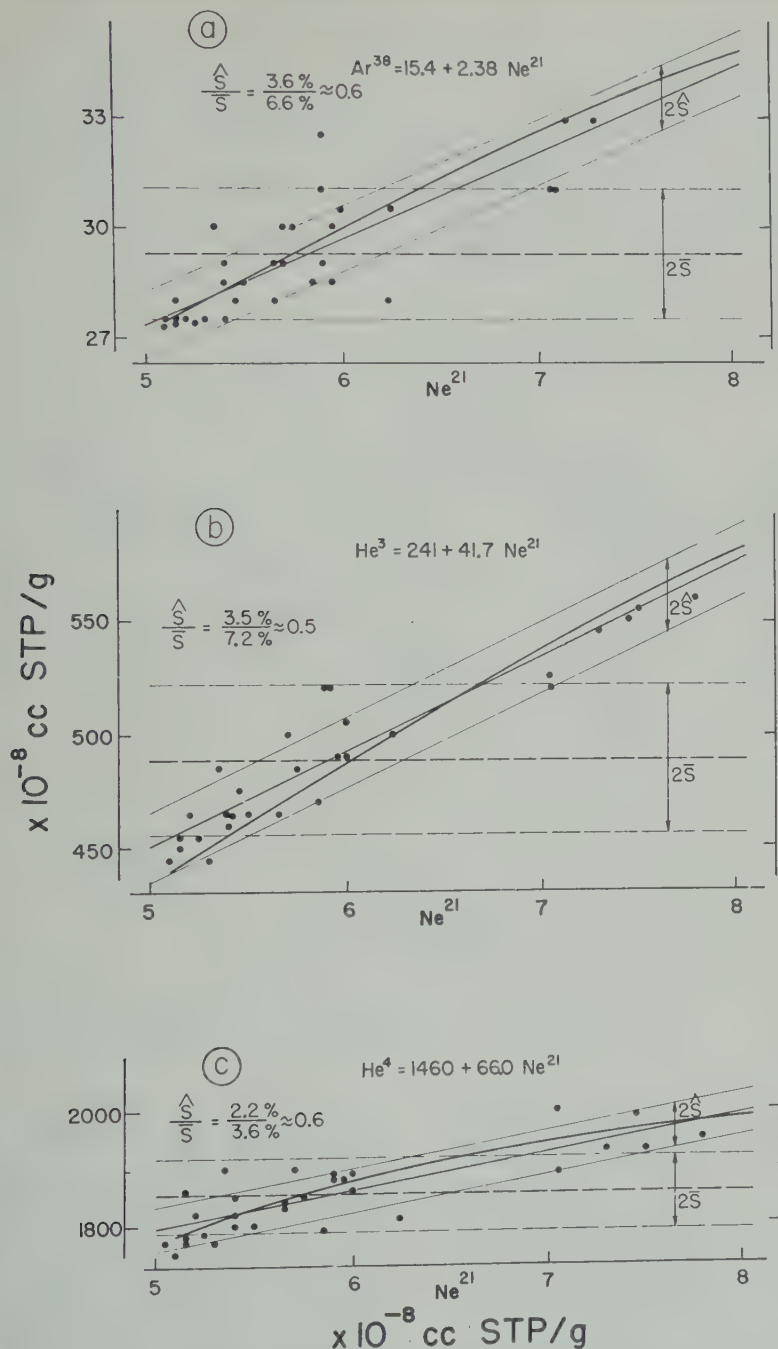


Fig. 3. Plots of  $Ar^{38}$ ,  $He^3$ , and  $He^4$  amounts vs.  $Ne^{21}$  amount in samples listed in Tables 1 and 2. Dashed lines show average values and associated standard deviations,  $\hat{S}$ . If a linear relation between  $Ar^{38}$  and  $Ne^{21}$  amounts is assumed, the heavier inclined line in (a),  $Ar^{38} = 15.4 + 2.38 Ne^{21}$ , gives correlation calculated by a least-squares analysis. Lighter parallel lines give corresponding standard deviation,  $S$ . Curves give correlation as computed from equation 3 discussed in section 3(c). In the figure, (b) and (c) give corresponding relations for  $He^3$  and  $He^4$ .

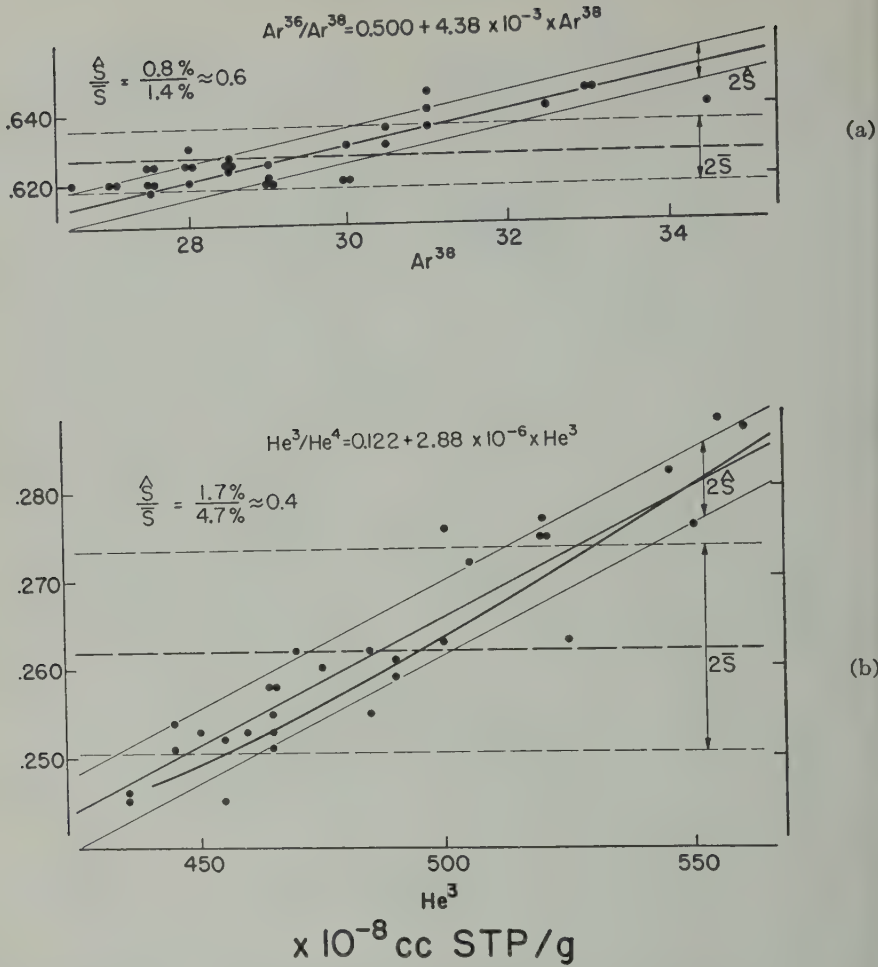


Fig. 4. Plots of  $\text{Ar}^{36}/\text{Ar}^{38}$  vs.  $\text{Ar}^{38}$  and  $\text{He}^3/\text{He}^4$  vs.  $\text{He}^3$  amount for all samples listed in Tables 1 and 2. Presentation and conventions as in Figure 3.

radial variations are then plotted as fractional increases from the center, the graphs obtained being shown in Figure 7.

(c) *Analysis of the results obtained in terms of a model for the production mechanism.* We will discuss the results in terms of the model proposed by Martin [1953] and Ebert and Wänke [1957] and employed by Hoffman and Nier [1958; 1959] in their discussion of depth studies of helium. In this model, if a unidirectional beam of primary cosmic rays is incident normally on a semi-infinite meteorite, and if it is assumed that:

- (a) the primary radiation is attenuated exponentially without any essential change in the energy distribution;
- (b) in the interaction of the primary particle with the target nucleus there are formed secondary particles of energy lower than primaries, but still sufficiently high to interact and also some high-energy particles (about 20 per cent according to Camerini, Davison, Fowler, Franzinetti, Muirhead, Lock, Perkins and Yekutieli [1951]);
- (c) wide-angle scattering processes can be neglected;
- (d) the production cross section for a particular nuclide is only a slowly varying function of

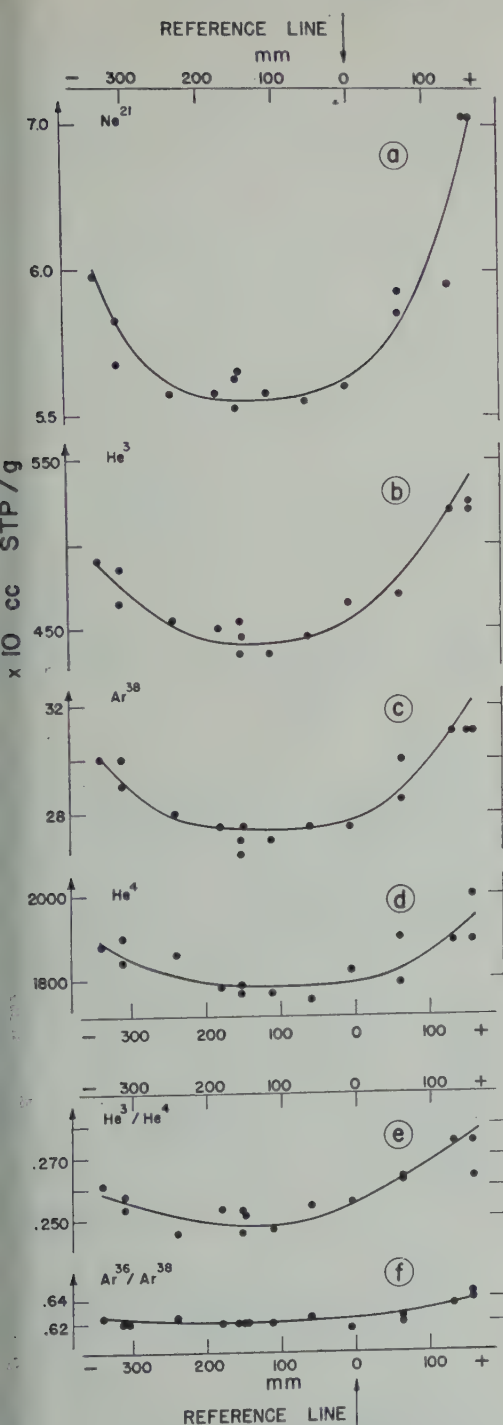


Fig. 5. Variation with position along bar F of  $\text{Ne}^{21}$ ,  $\text{He}^3$ ,  $\text{Ar}^{38}$ ,  $\text{He}^4$  (upper curves), and  $\text{He}^3/\text{He}^4$ ,  $\text{Ar}^{36}/\text{Ar}^{38}$  (lower curves).

energy and may be approximated by introducing two constant values, one for the primaries and one for the secondaries;

(e) tertiary and higher-order particles are counted as secondaries;

(f) radiation intensity was constant during the whole time of irradiation;

then one may write for the production of any nuclidic species  $i$ :

$$p_i = M_i \mu_p I_0 t e^{-\mu_a t}$$

$$+ m_i [\mu_s / (\mu_s - \mu_a)] S \mu_p I_0 t [e^{-\mu_a t} - e^{-\mu_s t}] \quad (2)$$

where

$M_i$ ,  $m_i$  = multiplicity of species  $i$ , i.e., the number of particles of species  $i$  produced per primary and secondary interaction, respectively;

$t$  = time during which irradiation took place;

$I_0$  = primary cosmic-ray flux;

$S$  = number of secondary particles emitted per primary interaction, assumed to be 2.9 [Martin, 1953];

and also

$$\mu_i = N \rho \sigma_i / A$$

where

$N$  = Avogadro's number;

$A$  = average atomic weight of meteorite material;

$\rho$  = density of meteorite material;

$j$  =  $a, p, s$ ;

$\sigma_a$  is effective absorption cross section of primary cosmic rays;  $\sigma_p$  is total interaction cross section of primary cosmic rays in a meteorite; and  $\sigma_s$  is secondary particle interaction cross section. It is assumed that  $\sigma_p = 1.2\sigma_a$ .

Since cosmic rays are actually omnidirectional rather than unidirectional, equation 2 must be integrated over all possible directions. Inasmuch as the rare gas production at any point due to cosmic rays coming from any particular direction depends upon the amount of meteorite material traversed by the beam, some assumption must be made about the shape of the meteorite. In the present, as in the earlier work [Hoffman and Nier, 1958], a spherical preatmospheric form is assumed.



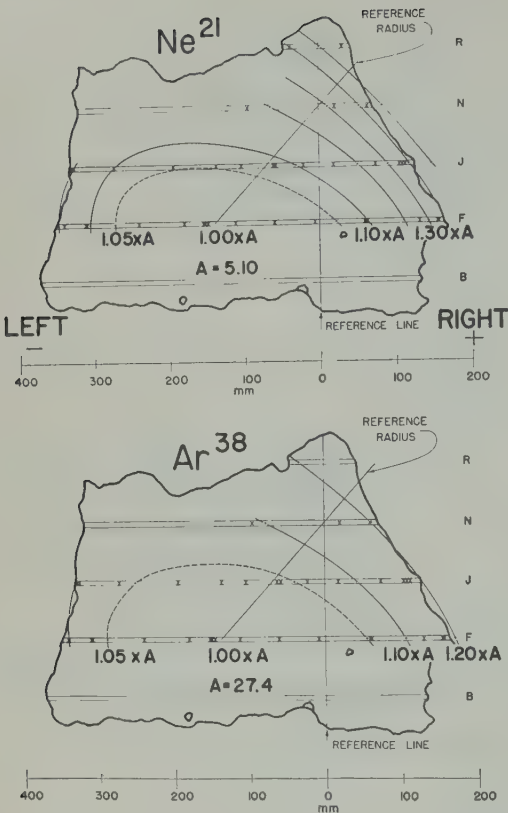


Fig. 6. Contours of equal concentration for  $\text{Ne}^{21}$  and  $\text{Ar}^{38}$ . Crosses (x) indicate points where analyses were made. The contours are not drawn as closed curves, since data were taken only for those regions which, as shown in the earlier investigations, had the lowest concentration and the biggest gradient in the concentration of the helium isotopes.

The integral form (2) may then be written as

$$P_i = A_i \left[ \int_0^{2\pi} \int_0^\pi e^{-k_a x} \sin \theta \, d\theta \, d\phi \right. \\ \left. - B_i \int_0^{2\pi} \int_0^\pi e^{-k_s x} \sin \theta \, d\theta \, d\phi \right] \quad (3)$$

where

- $R$  = preatmospheric radius of meteorite;
- $k_a = \mu_a R$ ;
- $k_s = \mu_s R$ ;
- $x = l/R$ ;

and  $A_i$  and  $B_i$  are given in terms of the coefficients appearing in (2).

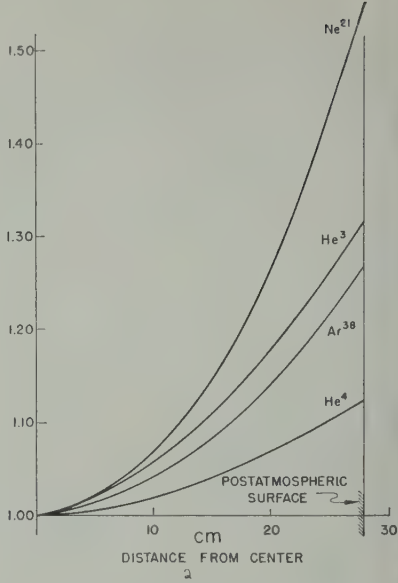


Fig. 7. Radial variation of  $\text{Ne}^{21}$ ,  $\text{He}^3$ ,  $\text{Ar}^{38}$ , and  $\text{He}^4$  determined from contour figures. Concentrations at center are 5.10, 440, 27.4, and  $1780 \times 10^8$  cc STP/g, respectively.

By adjusting the values of the parameters  $k_a$ ,  $k_s$ ,  $A_i$ ,  $B_i$  one attempts to fit equation 3 to the experimental data. It is found that the best fit is obtained if  $k_a$  and  $k_s$  have the values 1.10 and 1.92, respectively, and  $A_i$  and  $B_i$  have the values given in Table 3. It is interesting to note that a complete fit could be obtained only if  $B_{21} = 0$ . This is in agreement with the common assumption that the production cross section by secondaries in neon is negligibly small. Figure 8 shows the agreement between the experimental data and the values calculated from (3) using the parameters mentioned above. The agreement is also seen by referring to the curves plotted in Figures 3 and 4.

On the reference radius, the postatmospheric

TABLE 3. Values of  $A_i$  and  $B_i$  Giving Best Agreement between Experimental Data and Equation 3

| Parameter                  | <i>i</i> |      |      |      |
|----------------------------|----------|------|------|------|
|                            | 3        | 4    | 21   | 38   |
| $A_i \times 10^8$ cc STP/g | 296      | 1650 | 1.84 | 20.0 |
| $B_i$                      | 0.70     | 0.92 | 0.00 | 0.00 |

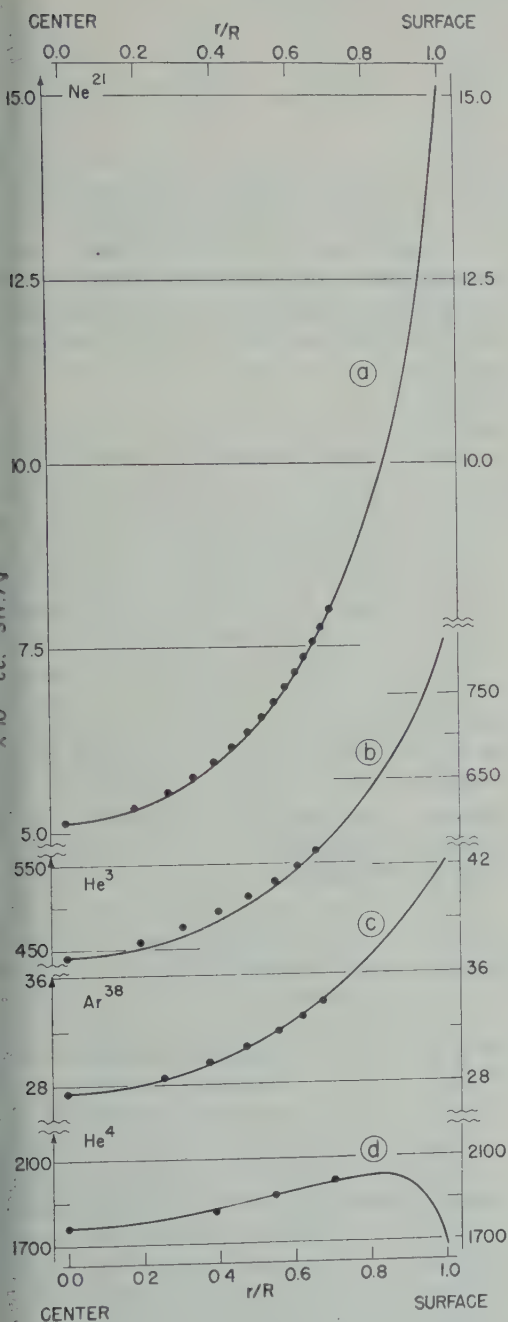


Fig. 8. Radial variation of (a)  $\text{Ne}^{21}$ , (b)  $\text{He}^3$ , (c)  $\text{Ar}^{38}$ , and (d)  $\text{He}^4$  as determined from equation (7). Dots are for experimental data (also given in Fig. 7) used in evaluating parameters in equation (7). Postatmospheric surface occurs at  $r/R = 0.7$ .

surface is at a distance of 28 cm from the center. From the fact that this point corresponds to a value of  $r/R = 0.7$  in Figure 8 we deduce that  $R = 40$  cm. This means, assuming the Grant to have been a spherical meteoroid, that its pre-atmospheric radius was 40 cm (and its mass approximately 2000 kg). This is in good agreement with the conclusions of the earlier work.

With the values of  $k_a$ ,  $k_s$  and  $R$  given here,  $\sigma_p = 1.2\sigma_a = 540$  mb and  $\sigma_s = 575$  mb. Hoffman and Nier [1958] found the same value for  $\sigma_p$  but a 25 per cent larger value for  $\sigma_s$ .

From the parameters given in Table 3, ratios of production cross sections can readily be calculated.<sup>1</sup> Table 4a gives values found in the earlier helium work together with corresponding values from the present investigation. The largest difference is 30 per cent. Considering the experimental difficulties and the differences in the averaging procedures the agreement can probably be considered satisfactory.

Table 4b gives a comparison between production ratios found in the present work and some preliminary measurements on a stainless-steel target bombarded with 6-bev protons. The target results are in agreement with values obtained by Bieri [1958] for a copper target also bombarded by 6-bev protons. In comparing meteorite and target results it should be noted that the meteorite, unlike the target, contains not only directly produced  $\text{He}^3$  but also some resulting from the decay of  $12.6$  y  $\text{H}^3$ . The ratio of production cross sections for  $\text{He}^3$  and  $\text{H}^3$  in the energy range under consideration here is not well established. A value of 0.8 for this ratio brings the  $\sigma_{p^3}/\sigma_{p^{21}}$  and  $\sigma_{p^3}/\sigma_{p^{38}}$  for Grant and

TABLE 4a. Ratios of Production Cross Sections in Helium

|                         | $\sigma_{p^3}/\sigma_{p^4}$ | $\sigma_s^3/\sigma_s^4$ | $\sigma_{p^3}/\sigma_s^3$ | $\sigma_{p^4}/\sigma_s^4$ |
|-------------------------|-----------------------------|-------------------------|---------------------------|---------------------------|
| Present work            | 0.67                        | 0.13                    | 4.8                       | 1.0                       |
| Hoffman and Nier [1958] | 0.50                        | 0.14                    | 5.0                       | 1.4                       |

<sup>1</sup> Since the half-lives of radioactive nuclides produced are short compared with the 'radiation age' of the meteorite, all the production cross sections of nuclides listed here include those of neighboring radioactive isobars.

TABLE 4b. Comparison between Production Ratios in Grant and a Stainless Steel Target Bombarded with 6-Bev Protons

|                        | $\sigma_{p^3}/\sigma_{p^4}$ | $\sigma_{p^3}/\sigma_{p^{21}}$ | $\sigma_{p^2}/\sigma_{p^{23}}$ | $\sigma_{p^4}/\sigma_{p^{21}}$ | $\sigma_{p^4}/\sigma_{p^{23}}$ | $\sigma_{p^{21}}/\sigma_{p^{23}}$ |
|------------------------|-----------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|-----------------------------------|
| Present work (Grant)   | 0.67                        | 48                             | 19                             | 72                             | 29                             | 0.40                              |
| Stainless-steel target | 0.17                        | 19                             | 10                             | 120                            | 54                             | 0.47                              |
|                        | 0.38*                       | 43*                            | 23*                            |                                |                                |                                   |

\* Target results corrected to include He<sup>3</sup> resulting from H<sup>3</sup> decay, assuming production ratio He<sup>3</sup>/H<sup>3</sup> =

target into agreement as indicated in Table 4b. Schaeffer and Zahringer [1959] indicated at 3 bev a value of 2.4 which appeared to be increasing with energy.

As shown in Table 4b, better agreement between meteorite and target would be obtained for ratios containing  $\sigma_p$ , if  $\sigma_p$  were larger than deduced from the present experimental data. A comparison with the results of Hoffman and Nier [1958] leads to the same conclusion.

In comparing meteorite and target results it must be remembered that the meteorite was exposed to particles having a continuous energy spectrum whereas the targets are bombarded by monoenergetic particles. Thus perhaps agreement should not be expected for all nuclides.

Table 4c gives the ratio of secondary to primary production cross sections as deduced from the parameters in equation 3. In our model the fact that these ratios differ from one another accounts for the variation of the nuclide ratios with depth in the meteorite.

The values of the parameters determined in the present work may be substituted in equation 2 to predict the variation with depth of a particular nuclide produced by a unidirectional beam. A computation of this kind was made for Ar<sup>36</sup>, and the result was compared with the measurements of Fireman and Zähringer [1957], who bombarded a long iron target with 6.2-bev protons and measured the Ar<sup>37</sup> produced; the experimentally measured amounts decrease faster

with depth than the calculated ones. Assuming that our model and parameters are correct, this could indicate that in long-target experiments some secondary particles are lost through scattering and hence do not contribute fully to production.

(d) Relation between rare gas amounts and nuclide ratios as a function of position in size of the Meteoroids. In his pioneering studies Martin [1953], using parameters deduced from emulsion work and theoretical data available at the time, showed how the helium content may vary with depth in spherical iron meteorites of various sizes. From parameters found in the present investigation it is possible to predict the depth variation of He<sup>3</sup>, He<sup>4</sup>, Ne<sup>21</sup>, and Figure 9 shows the calculated variations of the nuclides in spherical iron meteoroids having masses of 10<sup>3</sup>, 10<sup>4</sup>, 10<sup>5</sup>, and ∞ kg. The amounts of each isotope were normalized to their respective amounts to be expected in an infinitely small meteoroid.

The results of these calculations can be presented in several alternative ways in order to make evident the relation between the rare gas amounts and nuclide abundance ratios as a function of the position and the size of a meteoroid. Figure 10 shows how the He<sup>4</sup>/Ar<sup>36</sup> ratio varies with He<sup>3</sup>/Ne<sup>21</sup> in spherical meteoroids of various sizes. If cosmic rays have had a constant energy distribution in space and time, the ratios of nuclides as measured in a given sample should depend only upon the shielding which the sample had, i.e., the size of the original exposure body and the particular location of the sample in the body. Thus, for any sample from a spherical iron meteoroid, the plot of He<sup>4</sup>/Ar<sup>36</sup> versus He<sup>3</sup>/Ne<sup>21</sup> should lie on a curve belonging to the general family represented in Figure 9. If it does not, there must have been other sources of production of one or more of the nuclides

TABLE 4c. Ratios of Secondary to Primary Production Cross Sections as Computed from the Parameters  $k_a$ ,  $k_s$ ,  $A_i$ , and  $B_i$

| $\sigma_{s^3}/\sigma_{p^3}$ | $\sigma_{s^4}/\sigma_{p^4}$ | $\sigma_{s^{21}}/\sigma_{p^{21}}$ | $\sigma_{s^{23}}/\sigma_{p^{23}}$ |
|-----------------------------|-----------------------------|-----------------------------------|-----------------------------------|
| 0.21                        | 1.0                         | 0.00                              | 0.30                              |

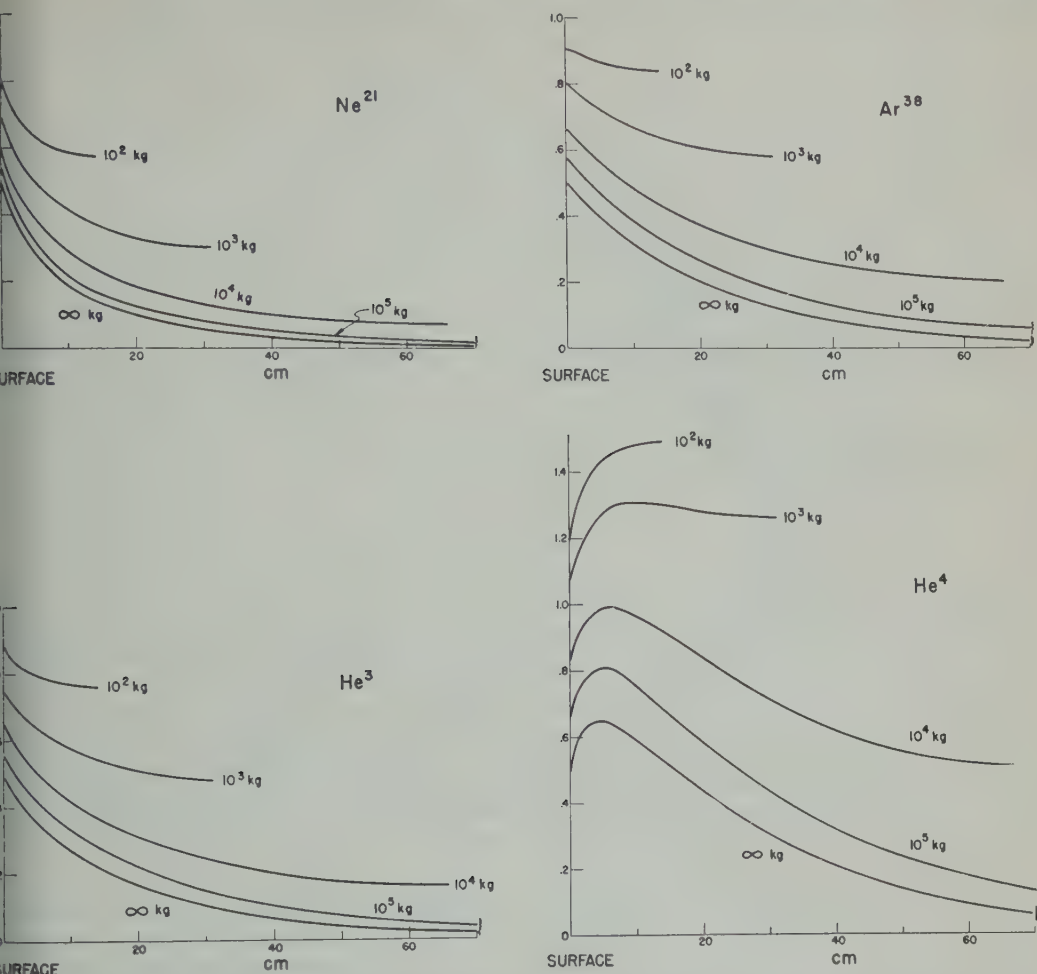


Fig. 9. Radial variation of Ne<sup>21</sup>, He<sup>3</sup>, Ar<sup>38</sup>, and He<sup>4</sup> for spherical meteoroids of different mass as predicted from equation 3. In each case the ordinate scale is chosen so that for an infinitesimally small meteoroid the amount would be unity.

for some alteration must have taken place in the meteoroid during the exposure. An example of a meteorite that would not fit Figure 10 is Washington County.

Another way to consider the relationships is shown in Figure 11. Here the He<sup>3</sup>/Ne<sup>21</sup> ratio is shown as a function of the Ar<sup>38</sup> content for spherical meteoroids of various sizes which received the same radiation dosage as the Grant meteorite upon which our numerical values for the parameters are based. The dashed curves give values that would be expected at the center and at the surfaces. On the assumption that the model and parameters are correct and all mete-

oroids are spherical and have the same radiation history as Grant, one should not expect to find points which lie outside of the area covered by the curves in Figure 11.

If a point lies on one of the curves in Figure 10, but outside the band of curves in Figure 11, it is concluded that the radiation dosage of the sample was different from that of Grant. Thus, if the point lies to the right, the dosage was greater than for Grant; if to the left, less.

Graphs such as Figure 11 can also be drawn for other combinations of nuclides and—if the model presented here is correct—should lead to consistent conclusions about the radiation dos-



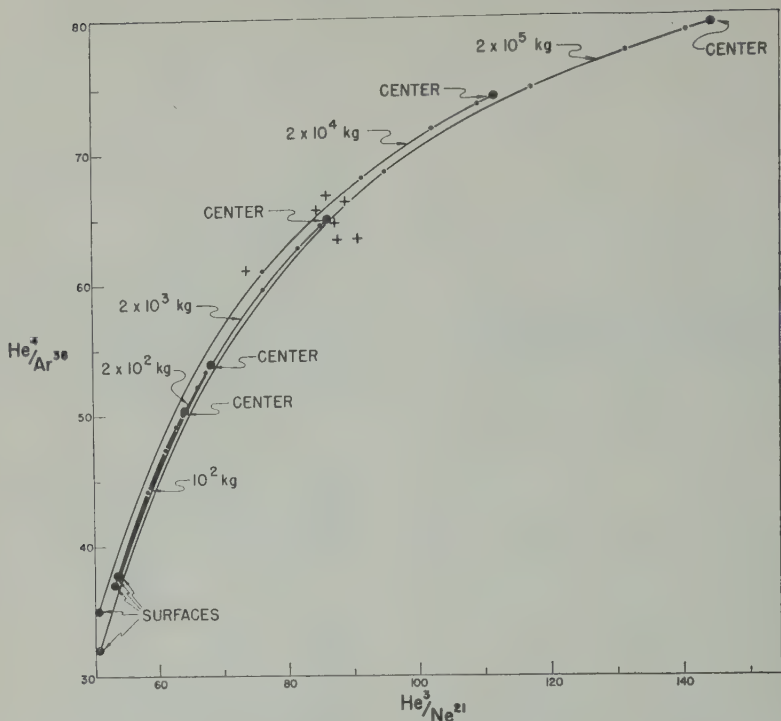


Fig. 10. Variation of the  $\text{He}^4/\text{Ar}^{38}$  with  $\text{He}^3/\text{Ne}^{21}$  in spherical iron meteoroids having masses of  $2 \times 10^2$ ,  $2 \times 10^3$ ,  $2 \times 10^4$ , and  $2 \times 10^5$  kg. Surface points lie at left end of curves, centers at right end. Points on curves correspond to depth with  $r/R$  values of 0.2, 0.4, 0.6, and 0.8, respectively. The large dots at the ends of the curves are for  $r/R$  values of 0.0 (center) and 1.0 (surface). Crosses are randomly chosen data (every other measurement) in Table 1.

age, the depth of the particular sample, and the original size of the meteoroid. Such alternative ways of examining the data serve to sharpen the interpretation.

From the measurements on Treysa published by Fechtig, Gentner, and Kistner [1960], we deduce by means of a graph analogous to Figure 11, correlating the isotopes  $\text{He}^3$ ,  $\text{He}^4$ , and  $\text{Ar}^{38}$ , that Treysa had a radiation age of 0.7 times that of Grant and a preatmospheric mass of  $\sim 1000$  kg ( $R \sim 30$  cm), and that the samples measured lay between the center and  $0.6R$ . These conclusions agree with those reached by the investigators.

The model, as presented so far, assumes that the irradiated body is spherical. Clearly, this may often be a poor approximation. To test the influence of shape, calculations were carried through for oblate and prolate ellipsoids of revolution. The ratios of the principal axes were chosen as 1 : 2 in both cases. Figure 12, *a* and *b*

(similar to Figs. 10 and 11), shows a comparison of the computations from three meteoroids of different shape each having a mass of 20000 kg. Note that the results are not very sensitive to shape, and hence, in view of the accuracy of the experimental procedure and other uncertainties in the entire analysis, the spherical approximation is reasonably good.

This study demonstrates that definite functional relationships exist correlating the relative amounts of cosmogenic nuclides with position and size of a meteoritic body. Such relationships are independent of whether the nuclides are stable or radioactive. Thus, in principle, the determination of a sufficient number of cosmogenic nuclides in a single sample may make it possible to determine the size and radiation history of the meteoroid as well as the position of the sample within the body. To exploit the possibilities fully, more quantitative data will be required than those presented in the present study.

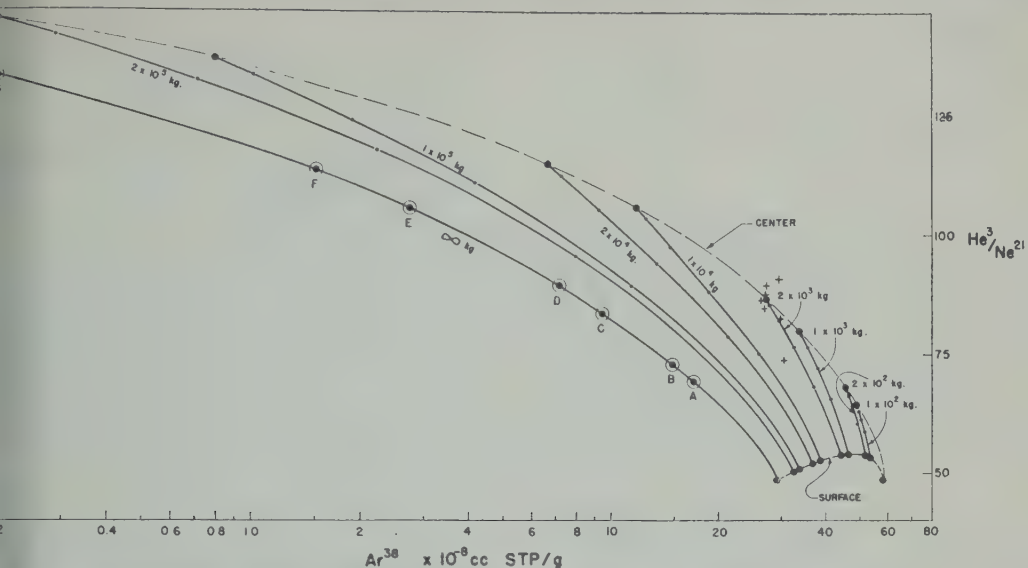


Fig. 11. Variation of  $\text{He}^3/\text{Ne}^{21}$  with  $\text{Ar}^{38}$  content as a function of position in spherical meteoroids of various sizes which receive the same radiation dosage as Grant. As in Figure 10, the points along the curves correspond to  $r/R$  values of 0, 0.2, 0.4, 0.6, 0.8, and 1.0, respectively. The encircled points A, B, C, D, E, F, and G shown on the infinitely large meteoroid curve correspond to values found at depths equal to the radii of the  $10^2$ ,  $2 \times 10^2$ ,  $10^3$ ,  $2 \times 10^3$ ,  $10^4$ ,  $2 \times 10^4$ , and  $10^5 \text{ kg}$  meteoroids, respectively.

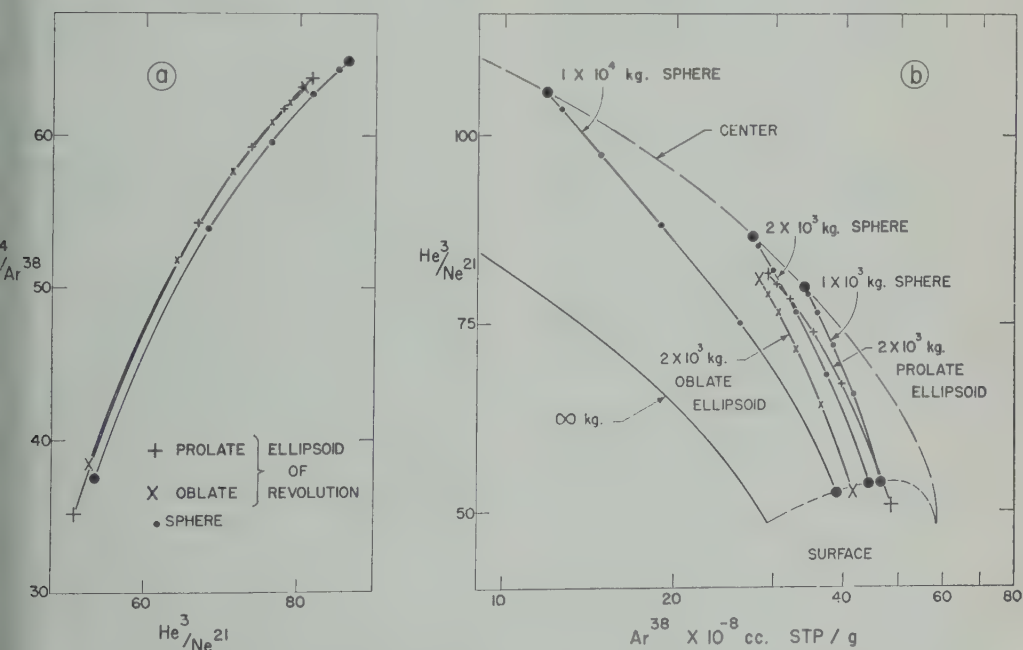


Fig. 12. Plots similar to Figures 10 and 11 for ellipsoids having masses of 2000 kg. Curves shown for axes of symmetry. For comparison purposes some data for spherical bodies are included. In Figures 12a all three bodies have identical mass, 2000 kg.

A valuable test of the model and the numerical parameters would be obtained from a depth study in a very large meteorite. Studies on meteorites of various sizes, including micro-meteorites and cosmic dust, would help develop the picture.

*Acknowledgments.* We are very much indebted to Mr. R. B. Thorness who, with his staff, constructed the apparatus employed in this work. Special thanks are due Dr. John H. Hoffman for providing numerical data needed in performing the integration of the production formula. The help of D. Maund, R. Oakes, and R. Stryk in carrying out parts of the experimental work is acknowledged.

The Grant meteorite material was provided by E. P. Henderson of the U. S. National Museum, Washington, D. C., and the targets by Dr. Walter Hartsough of the Lawrence Radiation Laboratory, Berkeley, California.

The research for this paper was supported by the joint program of the Office of Naval Research and the U. S. Atomic Energy Commission.

#### REFERENCES

- Bauer, C. A., The production of helium in meteorites by cosmic radiation, *Phys. Rev.* **72**, 354-355, 1947; **74**, 225-226, 1948; **74**, 501-502, 1948.
- Bieri, R., Rare gases in meteorite Carbo and a Cu target exposed to 6.2-bev protons, *Bull. Am. Phys. Soc.* [2], **3**, 221, 1958, and private communication.
- Camerini, U., J. H. Davies, P. H. Fowler, C. Franzinetti, H. Muirhead, W. O. Lock, D. H. Perkins, and G. Yekutieli, Nuclear transmutations produced by cosmic-ray particles of great energy, VI, Experimental results on meson production, *Phil. Mag.*, **42**, 1241-1276, 1951.
- Ebert, K. H., and H. Wänke, Ueber die Einwirkung der Höhenstrahlung auf Eisenmeteorite, *Z. Naturforsch.*, **12a**, 766-773, 1957.
- Fechtig, H., W. Gentner, und G. Kistner, Räumliche Verteilung der Edelgasisotope im Eisenmeteoriten Treysa, *Geochim. et Cosmochim. Acta*, **18**, 72-80, 1960.
- Fireman, E. L., Distribution of helium-3 in the Carbo meteorite, *Nature*, **181**, 1725, 1958.
- Fireman, E. L., The distribution of helium-3 in the Grant meteorite and a determination of the original mass, *Planetary and Space Sci.* **66-70**, 1959.
- Fireman, E. L., and J. Zähringer, Depth variation of tritium and argon-37 produced by high energy protons in iron, *Phys. Rev.*, **107**, 1169, 1957.
- Gerling, E. K., and L. K. Levskii, Cosmic radiation products in the Sikhote-Alin meteorite (English translation), *Doklady Akad. Nauk USSR*, **123**, 420, 1958.
- Hoffman, J. H., and A. O. Nier, Production of helium in iron meteorites by the action of cosmic rays, *Phys. Rev.*, **112**, 2112-2117, 1958.
- Hoffman, J. H., and A. O. Nier, The cosmogenic He<sup>3</sup> and He<sup>4</sup> distribution in the meteorite Carbo, *Geochim. et Cosmochim. Acta*, **17**, 32-36, 1953.
- Hoffman, J. H., and A. O. Nier, Cosmic-ray-produced helium in the Keen Mountain and Carbo meteorites, *J. Geophys. Research*, **65**, 1063-1068, 1960.
- Huntley, H. E., Production of helium by cosmic rays, *Nature*, **161**, 356, 1948.
- Martin, G. R., The origin of meteoritic helium and the age of meteorites, *Geochim. et Cosmochim. Acta*, **3**, 288-309, 1953.
- Paneth, F. A., P. Reasbeck, and K. I. Madsen, Helium-3 content and age of meteorite Carbo, *Geochim. et Cosmochim. Acta*, **2**, 300-303, 1952.
- Paneth, F. A., P. Reasbeck, and K. I. Madsen, Production by cosmic rays of helium-3 in meteorites, *Nature*, **172**, 200, 1953.
- Reasbeck, P., and K. J. Mayne, Cosmic radiation effects in meteorites, *Nature*, **176**, 733-734, 1955.
- Schaeffer, O. A., and D. E. Fisher, Cosmogenic noble gases in Washington County meteorites, *Nature*, **183**, 660-661, 1959.
- Schaeffer, O. A., and J. Zähringer, High sensitivity mass spectrometric measurements of stable helium and argon isotopes by high energy protons in iron, *Phys. Rev.*, **113**, 674-678, 1959.
- Stoener, R. W., and J. Zähringer, Potassium and argon ages of iron meteorites, *Geochim. et Cosmochim. Acta*, **15**, 40-50, 1958.
- Vinogradov, A. P., I. K. Zadorozhnyi, and I. P. Florenskii, Content of the inert gases in the Sikhote-Alin iron meteorite (English translation), *Geokhimiya*, **6**, 443-448, 1957.
- Wänke, H., and H. Hintenberger, Helium and Neon als Reaktionsprodukte der Höhenstrahlung in Eisenmeteoriten, *Z. Naturforsch.*, **9a**, 895-897, 1958.

(Manuscript received June 16, 1960.)

## The Origin of Meteorites

BRIAN MASON

*The American Museum of Natural History, New York*

**Abstract.** Current hypotheses generally interpret chondritic meteorites as being debris from a disrupted planet. It is suggested that these meteorites are not fragments of such a body, but have always been independent and individual objects. They can have been produced by the recrystallization of material now represented by the carbonaceous chondrites. The composition of carbonaceous chondrites is such that if they are heated above 600°C they would give a mixture of olivine, orthopyroxene, and nickel-iron similar to that of the chondritic meteorites. The friable and porous nature and the chondritic structure of these objects suggest recrystallization essentially in the solid state, as does the intimate admixture of nickel-iron and silicate. The other types of meteorites—achondrites, pallasites, and siderites—can be explained as fragments of a differentiated planetoid or planetoids formed by the aggregation of chondrites.

**Introduction.** The concept of meteorites as fragments of one or more disrupted planets has been so firmly fixed in the literature for so many years that it has become dogma. The purpose of this paper is to show that for the largest group of meteorites, the chondrites, this concept is unsatisfactory, and that they may never have formed part of a large body, or be the debris of such a body. The other groups of meteorites—irons, stony irons, and achondrites—do show features characteristic of differentiation in a comparatively large body and are probably products of disruption of a planet or planetoids.

In this connection it is illuminating to look at the statistics on meteorite abundance. Using the catalog of *Prior and Hey* [1953], we find that the numbers of falls of the different types recorded up to that time are: irons, 42 (6.6 per cent); stony irons, 12 (1.9 per cent); achondrites, 56 (8.8 per cent); and chondrites, 523 (82.7 per cent). It is unlikely that these percentages will be appreciably altered by future falls. The predominance of chondrites is extreme; a satisfactory theory of their origin is essential to our understanding of the origin and significance of all meteorites. It is first necessary to consider some of the features that are significantly related to the origin of chondrites—their chemical composition, their mineralogical composition, their texture and structure, and their ages.

**Chemical composition of chondrites.** At first glance a table of analyses of chondrites, such as

provided by *Urey and Craig* [1953], shows a great variation in the principal components:  $\text{SiO}_2$ ,  $\text{MgO}$ ,  $\text{FeO}$ , and  $\text{Fe}$ . However, it has been realized for many years that this variation is more apparent than real, being largely a reflection of the relative degree of oxidation of the iron. As a result,  $\text{FeO}$  ranges from essentially zero (0.21 per cent in Saint-Sauveur) to a maximum of about 26 per cent (25.43 per cent in Mokoia); concomitantly metallic  $\text{Fe}$  ranges from 25.73 per cent in Saint-Sauveur to 0.00 per cent in Mokoia. The nickel content of the metal phase shows an inverse relationship to the amount of this phase: the lower the metal content of a chondrite, the richer it is in nickel. These chemical regularities in the chondrites were codified some years ago by *Prior* [1916], in the form often known as Prior's rule: In any chondrite the richer in nickel is the nickel-iron, the richer in ferrous oxide are the magnesium silicates, or, in other words, the ratio of magnesia to ferrous oxide in the magnesium silicates varies directly with the ratio of iron to nickel in the nickel-iron.

It is clear from the preceding discussion that the  $\text{FeO}$  content is a significant figure in characterizing a chondrite. The dominant basic oxides in all chondrites are  $\text{MgO}$  and  $\text{FeO}$ . Figure 1 shows the numbers of chondrites having a specific  $\text{MgO}/(\text{MgO} + \text{FeO})$  ratio. This ratio ( $\times 100$ ) ranges from 60 to 100, with a large majority between 70 and 85. The lowest figure for the 'normal' chondrites (as distinguished from



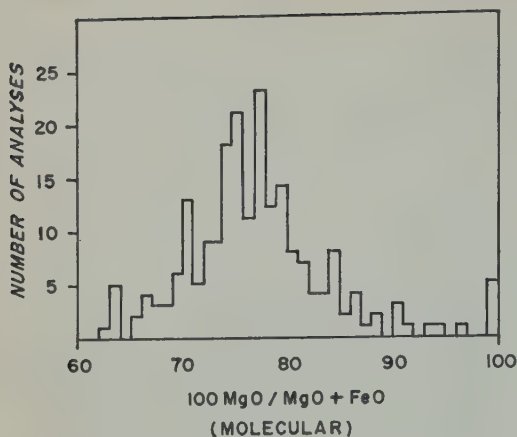


Fig. 1. Frequency distribution of analyses of chondritic meteorites in relation to molecular ratio  $100 \text{ MgO}/\text{MgO} + \text{FeO}$ .

the carbonaceous chondrites, which will be discussed later) is 63. For the carbonaceous chondrites this figure is between 50 and 60; precise figures are not easily derived from published analyses, many of which give all sulfur as  $\text{FeS}$ , whereas some is present as free sulfur and sulfate, and the  $\text{FeO}$  figure is then too low.

On the basis of this ratio Prior divided the chondrites into the enstatite chondrites, in which this ratio is 90 to 100; bronzite chondrites, ratio 80 to 90; and the hypersthene chondrites, ratio less than 80. This corresponds fairly closely to the actual presence of enstatite, bronzite, or hypersthene in these meteorites, although when the ratio is less than 66 the pyroxene is pigeonite rather than hypersthene.

Urey and Craig [1953] have distinguished two groups of chondrites on the basis of the total iron content. In their selected list of 94 analyses total iron ranged from 13.63 to 35.02 per cent; most analyses were in the 20 to 29 per cent range, and the frequency distribution showed peaks at 22 and 28 per cent. They used this relationship as a basis for arguing that the chondrites could not have originated from the breakup of a single large planet, and invoked an origin based on the breakup by collision of two asteroids of somewhat different composition.

*Mineralogical composition of chondrites.* Except for the carbonaceous chondrites, the mineralogical composition of the chondrites is extremely uniform, overlooking for the present the varia-

tion in composition of the individual mines. The principal constituents are olivine and pyroxene, which together make up 60 to 80 per cent by weight; the remaining minerals are nickel-iron, ranging up to 27 per cent, oligoclase or clinochlore, ranging up to 10 per cent, troilite about 6 per cent, and accessory amounts of chromite and a phosphate mineral, apatite or merrillite.

The relative amounts of the different minerals may be strongly dependent on the composition, as in olivine, pyroxene, and magnetite, or practically independent, as in oligoclase and troilite. This is shown schematically in Figure 2, where the weight percentages of the minerals are plotted against  $100 \text{ MgO}/\text{MgO} + \text{FeO}$ . As this ratio increases, both the olivine and pyroxene become progressively more magnesian in composition. Reference to Figure 1 shows that hypersthene chondrites are considerably more abundant than the other types; the Priory catalog lists as falls 129 hypersthene chondrites, 61 bronzite chondrites, and 15 enstatite chondrites.

In any individual chondrite, the olivine and pyroxene each have a fixed and uniform composition. This is clearly seen in X-ray diffraction records of these meteorites, which show sharp peaks for the individual reflections of these minerals. In addition, the relative com-

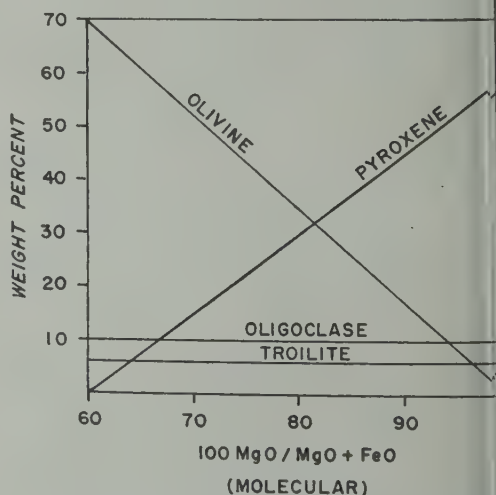


Fig. 2. Schematic representation of the mineralogical composition of chondritic meteorites in relation to molecular ratio  $100 \text{ MgO}/\text{MgO} + \text{FeO}$ .

tion of olivine and pyroxene shows a uniform relationship throughout: the olivine is slightly richer in iron than the coexisting pyroxene. This is consistent with equilibrium conditions in the system FeO-MgO-SiO<sub>2</sub> [Bowen and Schairer, 1935]. The uniformity of composition of the minerals in each individual chondrite, and their mutual relationship as exemplified by Prior's data, are powerful arguments against these meteorites' being chance aggregates of cosmic debris.

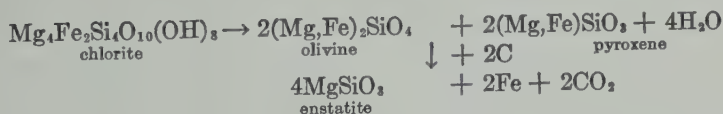
The carbonaceous chondrites form a distinct group mineralogically and chemically. Wüik [1956] has shown that in chemical composition they can be separated into three types, which he calls I, II, and III. Type III (the ornansites) consist essentially of olivine and have a comparatively small (<0.5 per cent) carbon content; they belong with the previously described chondrites rather than with the 'true' carbonaceous chondrites, types I and II, which contain 1 to 10 per cent free carbon, 10 to 20 per cent combined water (all other meteorites are essentially anhydrous), organic compounds, and free sulfur and sulfate compounds. Their mineralogy is not well known, because the finely divided carbon makes most if not all the material opaque to the microscope. Kvasha [1948] showed that the Boriskino carbonaceous chondrite is made up largely of chlorite. X-ray powder photographs of Alais, Cold Bokkeveld, Haripura, Mighei, Murray, Nawapali, Nogoya, and Santa Cruz show that these are also largely chlorite, or a related mineral with the structure of serpentine; X-ray powder photographs of Orgueil, Ivuna, and Kaba show lines of magnetite only. The carbon in the

aminated. They show none of the prominent characteristics of other meteorites: they contain no nickel-iron; they are of low density (about 2.5); they look like a black friable clay. Moreover, they would not survive even mild weathering; some of them break up rapidly in water as a result of the leaching of soluble compounds.

In spite of the extreme mineralogical contrasts between carbonaceous chondrites and 'normal' chondrites there is a close chemical relationship between the two groups. This has been brought out by Wüik [1956], who recalculated analyses of carbonaceous and other chondrites on a water, carbon, oxygen, and sulfur-free basis. When this is done the elemental composition of the carbonaceous chondrites matches closely that of the olivine-pyroxene chondrites.

The different groups of chondrites thus show a remarkable sequential relationship in their chemical and mineralogical composition. The carbonaceous chondrites contain no free iron and have a 100 MgO/MgO+FeO ratio of 60 or less. The ornansites (often grouped with the carbonaceous chondrites since they contain about 0.5 per cent free carbon) consist essentially of olivine with little pyroxene or metal, and the ratio is 60 to 65. With increasing pyroxene and decreasing olivine we pass to the hypersthene chondrites (ratio 65 to 80), bronzite chondrites (ratio 80 to 90), and enstatite chondrites (ratio 90 to 100).

This can readily be interpreted in terms of the thermal decomposition of chlorite, which takes place above 600°C, coupled with the reduction of part or all of the combined iron to metal by carbon, according to the equation



carbonaceous chondrites is amorphous to X rays, and it appears that much other material must be amorphous also. An analysis of Orgueil shows 2 per cent SiO<sub>2</sub>, yet no silicate mineral appears in the powder photograph.

Carbonaceous chondrites are rare. The Prior's catalog lists 19, all falls. The reason for their being all falls is obvious when they are ex-

According to this formulation the thermal decomposition of chlorite gives equal amounts of olivine and pyroxene; however, carbonaceous chondrites commonly contain excess iron oxide in the form of magnetite, and under these circumstances an olivine chondrite of the ornansite type would be formed. This reaction has been confirmed for Cold Bokkeveld and Murray by heating samples of these carbonaceous chon-

drites in vacuum; the product consisted of olivine, possibly accompanied by a little nickel-iron in the Cold Bokkeveld sample.

Judging by the relative abundance of the meteorite types, if the 'normal' chondrites were formed from the carbonaceous chondrites in this way, the amount of reduction commonly resulted in the formation of a hypersthene chondrite, and the more reduced types (bronzite and enstatite chondrites) are progressively less abundant. As was mentioned above, the amount of carbon in the carbonaceous chondrites varies considerably, from about 1 to 5 per cent; this is the right order of magnitude to produce any degree of reduction from the hypersthene chondrites to the pure enstatite chondrites.

*Texture and structure of chondrites.* Chondrites are so named because of the presence of chondrules, which are small ( $\sim 1$ -mm diameter) spheroidal aggregates of one or more silicates, generally olivine and/or pyroxene. In form, manner of crystallization, texture, and structure, they show little resemblance to any spheroidal bodies observed in terrestrial rocks. Their origin has been a subject of controversy for over a century.

Roy [1957] has summarized the theories of origin of chondrules as follows: (1) chondrules are fused drops of 'fiery rain'; (2) chondrules are fragments of pre-existing meteorites, which have become rounded by oscillation and attrition; (3) chondrules are products of a special phase of magmatic segregation, formed in place as a result of rapid, arrested crystallization in a molten mass; (4) chondrules originated from a dispersal of a silicate melt in a hot atmosphere, the resultant drops crystallizing from the outside inward; (5) chondrules are metamorphosed garnets—garnets converted to enstatite; (6) chondrules were produced by the cooling of liquid silicates, which fell as a molten rain during the collision of a small asteroid with a larger one. Roy favors the formation of chondrules in place, as products of separation from a silicate melt, followed often by subsequent deformation and metamorphism. Ringwood [1959], who has recently discussed the origin of chondrules in considerable detail, shows inadequacies in most of these hypotheses. He adduces evidence for the formation of chondrules at moderate temperatures, well below the normal melting points of the minerals composing them.

The chondritic structure of olivine-pyroxene chondrites can be a feature inherited from the carbonaceous chondrites. The carbonaceous chondrites contain chondrules of chlorite. Crystallization of these chondrules without melting would produce the olivine and pyroxene chondrules of the common chondrites. In connection it is worth remarking that enstatite chondrites commonly show the chondritic structure much less prominently than the hypersthene and bronzite chondrites. This is understandable if they represent material that has been subjected to the greatest degree of reduction and recrystallization.

Of course, interpreting the chondritic structure as an inherited feature from the carbonaceous chondrites does not solve the origin of chondrules themselves. It is conceivable that chlorite chondrules in the carbonaceous chondrites are the product of crystallization of amorphous material at comparatively low temperatures, analogous to the growth of crystalline spherules in a devitrifying glass. Some carbonaceous chondrites, such as Orgueil, contain or no crystalline silicate, the silicate material evidently being present largely in the amorphous state. It is noteworthy that Orgueil contains visible chondrules. Crystallization of the amorphous silicate material at low temperatures could well produce spherules, i.e., chondruleless chlorite.

The texture and structure of the chondritic meteorites indicate that they have not originated in any body with a considerable gravitational field. Many of them are very porous and friable to a degree that they can be crumbled in the hand, indicating that they were not consolidated under pressure (this is particularly characteristic of the falls; finds are frequently compacted through cementation by secondary iron oxides formed by terrestrial weathering). The nickel-iron in the chondrites is intimately mixed with the low-density silicate minerals; it shows no sign of gravitational segregation. This relationship is consistent with the formation in place of the metal by reduction from the silicates. The belief that the nickel-iron in chondrites formed in place by the reduction of compounds of iron is not a novel one; it was expressed as long ago as 1878 by Nordenskiöld. It has been repeated by more recent observers. Urey and Mayeda [1959] point out that many



metallic particles in chondrites are distorted and their borders are broken and twisted; they conclude that this distortion and breaking argues for an origin in a 'primary' body which was broken up and the fragments incorporated in the chondrites. However, if the particles of metal in chondrites are fragments from a disrupted body, why is their composition directly related to the bulk composition of the meteorite? This indicates a chemical equilibrium established within each individual chondrite, not a chance aggregate of fragments. The distortion of the metallic particles described by Urey and Heyeda could be a later feature, produced by impact between chondrites or by impact of the chondrite with the earth.

*Ages of chondrites.* In recent years a great deal of research has been carried out on the history of meteorites as revealed by different methods of dating. Urey [1959] has compiled the potassium-argon ages for twelve chondrites; these ages range from  $0.65 \times 10^9$  years to  $4.5 \times 10^9$  years. These values might be interpreted as an indication of different times of crystallization for different chondrites. However, it has also been suggested that the low ages of some chondrites are spurious, resulting from diffusive loss of argon. Further data are evidently required before it can be decided whether all chondrites crystallized in a single brief period or at widely different times.

Cosmic-ray-exposure ages have been measured on a number of meteorites. The figures for fourteen chondrites as given by Urey [1959] range from <2 to 90 million years. These meteorites as collected have thus been exposed to cosmic rays for periods of less than  $10^8$  years, and this result has been used to support the theory that chondrites were produced by the breakup of a planet or similar body. Fireman and DeFelice [1960], who have recently discussed the possible interpretation of the cosmic-ray ages, find the planetary breakup theory unsatisfactory. They prefer a theory that interprets the cosmic-ray ages in terms of erosion of the meteorites in extraterrestrial space by the action of dust, gas, ions, and electrons. As a result of this erosion the measured cosmic-ray age is less than the interval since the meteorite was first exposed to cosmic rays. Such a theory indicates that the meteorites have been somewhat more massive in outer space but does not require

that they have been shielded within a large body of planetary or lunar size.

*Theories of meteorite origin.* Many theories have been proposed for the origin of meteorites. H. A. Newton, as long ago as 1886, summarized the numerous theories as follows: 'They came from the moon; they came from the earth's volcanoes; they came from the sun; they came from Jupiter and the other planets; they came from some destroyed planet; they came from comets; they came from the nebulous mass from which the solar system has grown; they came from the fixed stars; they came from the depths of space' [Newton, 1886]. At the present time the field of hypotheses has been considerably narrowed, and most authorities prefer an origin based on a destroyed planet or other relatively large body. Urey has studied this problem intensively in recent years, and in his Hugo Müller lecture [Urey, 1958] summarized his views as follows: 'It is proposed that the observations are in accord with the following series of events. (1) A disc nebula at temperatures so low that hydrogen is partly condensed to the solid state existed, though no process for its formation is discussed. (2) This broke up into masses of lunar size together with the quota of cosmic gases. (3) These objects collided later to produce very fine solid material, and part of this material and the gases were lost to space owing to particle radiation from the sun. (4) Subsequently the residue accumulated into the planets and the immediate parents of the meteorites.'

Some of the facts regarding the chondrites are inconsistent with their origin either as massive fragments of a disrupted planet or as aggregated dust from such a disruption. The mineralogy of the typical carbonaceous chondrites (excluding the ornansites) shows that they have never been above  $600^\circ\text{C}$  and probably have always been at much lower temperatures. The texture and structure of the noncarbonaceous chondrites show that they are not fragments of a disrupted planet, for then gravitational segregation of silicate and metal should have taken place. Their mineralogy, especially the uniformity of composition of olivine and pyroxene in each individual meteorite, shows that they are not chance aggregates of dust but individual chemical systems in phase equilibrium. Their compositions are consistent with an origin by



recrystallization of the material of carbonaceous chondrites. Such an origin does not require that they were ever part of a body or bodies of planetary or lunar dimensions, and their texture, their structure, and their mineralogy speak directly against such an origin.

*Acknowledgments.* I am indebted to Dr. Felix Chayes and Dr. E. L. Fireman for critical reading of the manuscript and suggestions for its improvement, and to Dr. E. L. Fireman for heating samples of Murray and Cold Bokkeveld in a vacuum furnace.

#### REFERENCES

- Bowen, N. L., and J. F. Schairer, The system  $MgO-FeO-SiO_2$ , *Am. J. Sci.*, **29**, 151-217, 1935.
- Fireman, E. L., and J. DeFelice, Argon 39 and tritium in meteorites, *Geochim. et Cosmochim. Acta*, **18**, 183-192, 1960.
- Kvasha, L. G., Investigation of the stony meteorite Staroye Boriskino, *Meteoritika*, **4**, 83-96, 1948.
- Newton, H. A., Meteors, meteorites, and shooting stars, *Popular Sci. Monthly*, **29**, 733-747, 1886.
- Nordenskiöld, A. E., Trenne markeliga eldmeteorer, sedda i Sverige under åren 1876 och 1877, *Geol. Fören. i Stockholm Förh.*, **4**, 45-61, 1916.
- Prior, G. T., On the genetic relationship and classification of meteorites, *Mineral. Mag.*, **18**, 27, 1916.
- Prior, G. T., and M. H. Hey, *Catalog of Meteorites*, 2d ed., British Museum, London, 1931.
- Ringwood, A. E., On the chemical evolution and densities of the planets, *Geochim. et Cosmochim. Acta*, **15**, 257-283, 1959.
- Roy, S. K., The problems of the origin and structure of chondrules in stony meteorites, *Fieldeol.*, **10**, 383-396, 1957.
- Urey, H. C., The early history of the solar system as indicated by the meteorites, *J. Chem. Soc.*, 67-78, 1958.
- Urey, H. C., Primary and secondary objects, *Geophys. Research*, **64**, 1721-1737, 1959.
- Urey, H. C., and H. Craig, The composition of stony meteorites and the origin of the meteorites, *Geochim. et Cosmochim. Acta*, **4**, 36-82, 1939.
- Urey, H. C., and T. Mayeda, The metallic part of some chondrites, *Geochim. et Cosmochim. Acta*, **17**, 113-124, 1959.
- Wiik, H. B., The chemical composition of stony meteorites, *Geochim. et Cosmochim. Acta*, **9**, 279-289, 1956.

(Manuscript received May 5, 1960; revised June 21, 1960.)

# Black, Magnetic Spherules in Sediments

W. D. CROZIER

*New Mexico Institute of Mining and Technology  
Socorro, New Mexico*

**Abstract.** Specimens of sedimentary materials from formations distributed over the sequence from Recent to Ordovician have been processed for recovery of magnetic particles. Magnetic particles were found in all specimens, and among them were black spherules in numbers ranging from 28 to 240 per gram of sediment. These are apparently identical to and have the same size distribution as the black, magnetic spherules currently being collected from the atmosphere and to the magnetic spherules of supposed meteoritic origin that for many years have been reported in deep sea sediments.

Two specimens of deep sea sediment were among those studied. They were found to contain larger spherules in numbers of the same order as those reported by Laevastu and Mellis, but spherules smaller than  $40\ \mu$  in diameter were much more abundant. A large and, thus far, unresolved discrepancy appears to exist between the annual infall of spherules calculated for the deep sea sediments and that calculated for the San Agustin playa deposits and for the present atmospheric deposit.

## INTRODUCTION

From time to time, apparently beginning with Murray [1876], small, black, magnetic, spherulitic particles, recognized as being of probable meteoritic origin, have been described as occurring in deep sea sediments. These reports culminate, perhaps, in several papers dealing with black, magnetic spherules in cores obtained during the Swedish Deep-Sea Expedition of 1947-48. [Laevastu and Mellis, 1955; Pettersson and Fredriksson, 1958]. Laevastu and Mellis reported recovery of up to 2000 spherules per gram of red clay in a core from the sea bottom in the central Pacific ( $7^{\circ}38'S$ ,  $152^{\circ}W$ ); these sediments were of ages up to an estimated 7 million years. It long has been supposed that the magnetic spherules should be recoverable from sediments other than those laid down in the deep seas, but there seems to have been practically no definite success in such recovery until very recently. The author [Crozier, 1957] reported finding considerable numbers of these objects in playa deposits on the San Agustin Plains, in New Mexico, and subsequently [Crozier, 1958] reported finding them in other sediments. Erskine [1959] reported their recovery from the Todilto gypsum, and Skolnick [1959] reported them in California in well cuttings and cores of Miocene and younger ages.

In this paper the results of a modest number of 'spot checks' in sediments of ages running back to the Ordovician are reported.

## METHOD OF SPHERULE RECOVERY

In outline, the steps of the recovery procedure include (1) dispersion, to get as much as possible of the rock or unconsolidated material into suspension or solution; (2) elutriation, for removal of clay and fine silt; (3) a special, delicate, magnetic separation; and (4) microscopic study, including the mounting of the magnetic fraction on membrane filter disks, followed by counting and classification of the black spherules.

**Dispersion.** With some clays and shales, satisfactory dispersion can be effected by simply stirring them up in a 0.01-N solution of Calgon (sodium hexametaphosphate). The dispersion is usually not complete at first, and additional dispersion takes place during succeeding repeated cycles of elutriation in the Calgon solution. In some specimens containing salts, preliminary treatment with 1-N HCl is used, possibly with a slight risk of loss or alteration of the magnetic spherules. In some control experiments, however, it appeared that 12 to 24 hours' exposure to 1-N HCl did not significantly reduce the number of spherules counted. Exposure to the

acid for a week resulted in a considerable reduction in the number, however. In the case of some shales, mild methods have failed to disperse the sample, and further processing has not been attempted.

The limestone and dolomites are treated with HCl until all the carbonate is dissolved; the residue is then treated in the same manner as the clays and shales.

*Elutriation.* The suspensions are shaken thoroughly in a jar with a 10-cm depth of liquid and allowed to settle  $2\frac{1}{2}$  hours. This is sufficient to permit all spherules larger than  $2\ \mu$  in diameter and of density greater than  $5\ \text{g/cm}^3$  to be deposited on the bottom. The suspension is then decanted very slowly through a siphon, to avoid picking up sediment from the bottom. After decanting, new Calgon solution is added to the sediment, and the cycle is repeated; from 10 to 20 cycles usually are required for satisfactory elimination of clay and fine silt. The jar is kept enclosed during the settling period to minimize convection. Precautions are taken to prevent contamination of spherules from the reagents or from the laboratory atmosphere; all liquids added to the sample have been filtered through a Millipore membrane filter, air drawn into the jar during decanting passes through a tube containing a tight plug of glass wool, and the jar is refilled through the siphon tube.

The elutriation process undoubtedly is accompanied by some loss of the magnetic spherules, since it is known that the net density of a small portion of the spherules is considerably reduced by enclosed gas bubbles. Some studies of the supernatant liquid indicated that the loss probably does not exceed 20 per cent; whenever losses of this order become significant, the matter will bear further investigation.

*Magnetic separation.* After the final decanting, a magnetic separation is performed on the remaining particles in a specially designed apparatus. For recovery of very small magnetic particles from a fluid, a separation technique must be used which not only provides a strong magnetic field, with a large gradient in the direction of the desired particle motion, but also provides short paths for the particle motion and allows sufficient time for the traversal of the paths. This is a consequence of the fact, apparently not always fully appreciated by

workers in the field, that, given the magnetic field strength, the gradient of field strength, susceptibility of the particles, and the viscosity of the fluid, the translational velocity of the particles decreases about as the second power of particle diameter. In the apparatus used, the requirements are met by letting all the particles settle to the flat bottom of an aluminum cup from a suspension depth of 2 to 3 cm, then, by means of an appropriate mechanical assembly slowly passing over the bottom of the cup the end of a strong Alnico bar magnet enclosed in a thin (0.25-mm) brass or copper sheath. A large gradient appears at the intersection of the flat end with the cylindrical side surface of the magnet. Usually it is possible to bring the spacing between the sheath and the bottom of the cup down to 0.6 mm or less. The mechanical assembly slowly rotates the cup under the magnet at the same time that the magnet traverses the radius (4 cm) of the cup. About 15 minutes of time are allowed for complete coverage of the bottom.

After the traverse, the magnet and sheath assembly is slowly raised by a screw to a distance of about 1.5 cm from the bottom of the cup, and a small glass dipper is slipped under its end. Then the magnet, sheath, and dipper, with the water in the dipper, are transferred together to a small jar of clean water and immersed as they were in the aluminum cup. The magnet is withdrawn quickly from the sheath, the viscosity of the water restraining the magnetic particles from rolling up the sheath, and the particles are washed from the sheath and dipper into the small jar.

The particles remaining in the aluminum cup are stirred vigorously to break up 'traps' which may have prevented recovery of some magnetic particles in the first cycle, and the whole process is repeated. From four to a dozen such cycles, the number depending upon the magnet spacing and the amount of nonmagnetic material present, make it possible to transfer most of the magnetic particles to the clean water, with only a moderate contamination by nonmagnetic particles.

*Microscopy.* The water containing the magnetic particles is passed through a 2-cm Millipore filter disk, leaving the particles on the surface of the disk. If there is a considerable amount of magnetic material, it may be divided



TABLE 1. Content of Black, Magnetic Spherules in Various Sedimentary Rocks

| Formation       | Age           | Sample Size, g | Number of Magnetic Spherules per Gram |                   |                    |                    |           | Total |
|-----------------|---------------|----------------|---------------------------------------|-------------------|--------------------|--------------------|-----------|-------|
|                 |               |                | <5 $\mu$                              | 5 $\mu$ -10 $\mu$ | 10 $\mu$ -15 $\mu$ | 15 $\mu$ -20 $\mu$ | >20 $\mu$ |       |
| ica shale       | Ordovician    | 2.3            | 53                                    | 33                | 9                  | 1.7                | 1.7       | 98    |
| ayman shale     | U. Silurian   | 2.1            | 54                                    | 14                | 3                  | 0.9                | 0.9       | 73    |
| ekport dolomite | M. Silurian   | 2.4            | 20                                    | 7                 | 4                  | 0.4                | ...       | 32    |
| ale             | Silurian      | 2.0            | 127                                   | 95                | 16                 | 1.5                | 1.0       | 241   |
| adera limestone | Pennsylvanian | 3.3            | 10                                    | 5                 | 1                  | ...                | ...       | 16    |
| ale (maroon)    | Triassic      | 2.4            | 25                                    | 16                | 2                  | 0.8                | 1.2       | 45    |
| ale (gray)      | Triassic      | 2.0            | 5                                     | 3                 | 2                  | 0.5                | ...       | 10    |

among several disks. The Millipore disks are mounted on microscope slides, dried, treated with immersion oil to make them transparent and covered with glass slips.

The opaque, smooth, black spherules in the mounted magnetic fraction are counted under dark-field illumination, with X200 magnification, and the diameter of each spherule is measured with an eyepiece reticle scale. Identification is sometimes aided by the use of oblique top illumination, under which the magnetic spherules exhibit a metallic luster.

Particles with silhouettes departing from true circularity, even if only slightly, are not counted; neither are those with appreciably rough surfaces. In spherules examined in air, it is sometimes possible with critical oblique illumination to see a suggestion of mosaic markings; these do not cause departure from smooth sphericity, and they are not visible ordinarily when the spherules are mounted in oil.

## RESULTS

Because of the many stages of processing and tedious microscope work involved, the study has been limited to a fairly small number of sediment samples. Also, in general, the sample size has had to be held to the range of 2 to 4 grams of rock, to avoid overloading the magnetic separator. Samples, as far as possible, were selected from chips from the interior of hand specimens, to avoid surface contamination. Veinings were avoided, except in a few cases where very fine veining made this impossible.

All the sediments were found to contain many magnetic particles of various shapes and sizes, in addition to the number in spherule form.

Since the magnetic fraction must be spread out rather thinly on the Millipore disks to enable the microscopists to scan particle by particle, this factor also contributed to the limitation of practicable sample size.

**Rocks.** Table 1 shows the results of magnetic-spherule counts for a selection of available rock samples, covering an age range from Ordovician to Triassic. The first four samples in the table were collected in New York; the last three are from the Carthage, N. Mex., area. Table 2 shows the results of a slightly more systematic study of some Upper Cretaceous shales, most of them from the Mancos formation, also collected in the Carthage, N. Mex., area. The Wapiabi shale is from a core taken 5740 feet below the surface in the Texaco Glendale well, in Canada, obtained through the courtesy of Professor R. S. Taylor. The specimen marked "Datil Mts." is from an outcrop in these mountains (N. Mex.) and probably is to be classed as a Mancos shale. The last specimen is from the Big Bend area, in Texas.

Since the rock samples are so small, the statistical significance which can be attached to the absolute numbers of magnetic spherules counted in this study is somewhat limited. An additional limitation is to be found in the fact that no two microscopists count the same number of spherules in a given slide, the variations sometimes being rather large. Perhaps the most significant result is that no sedimentary material has yet been found that is completely devoid of the spherules.

In spite of the low statistical significance of the spherule counts, there are some features of the results in Tables 1 and 2 that deserve comment. For example, of the two specimens of



TABLE 2. Content of Black, Magnetic Spherules in Some Cretaceous Shales

| Formation    | Depth<br>below<br>Top, ft | Sample<br>Size, g | Number of Magnetic Spherules per Gram |                  |                   |                   |            |
|--------------|---------------------------|-------------------|---------------------------------------|------------------|-------------------|-------------------|------------|
|              |                           |                   | $<5\ \mu$                             | $5\ \mu-10\ \mu$ | $10\ \mu-15\ \mu$ | $15\ \mu-20\ \mu$ | $>20\ \mu$ |
| Lower Mancos | 150                       | 2.0               | 39                                    | 21               | 4                 | 1.5               | 1.5        |
|              | 125                       | 2.0               | 67                                    | 20               | 5                 | 1.0               | 0.5        |
|              | 60                        | 2.0               | 38                                    | 16               | 3                 | 0.5               | 1.5        |
|              | 25                        | 2.0               | 39                                    | 16               | 2                 | 0.5               | 0.5        |
| Upper Mancos | 40                        | 2.1               | 41                                    | 28               | 8                 | 0.5               | 2.0        |
|              | 20                        | 2.0               | 57                                    | 22               | 6                 | 0.5               | 0.5        |
|              | 15                        | 2.1               | 10                                    | 7                | 3                 | 1.4               | 1.4        |
|              | 10                        | 2.0               | 111                                   | 33               | 8                 | 3.0               | 0.5        |
|              | 10                        | 3.6               | 79                                    | 42               | 6                 | 0.3               | 1.1        |
|              | 5                         | 2.1               | 15                                    | 10               | 4                 | ...               | ...        |
| Wapiabi      | ...                       | 2.4               | 28                                    | 21               | 6                 | 1.7               | ...        |
| "Datil Mts." | ...                       | 1.8               | 52                                    | 37               | 4                 | 1.6               | ...        |
| "Big Bend"   | ...                       | 4.0               | 15                                    | 7                | 4                 | 0.5               | 0.3        |

Triassic shale in Table 1, collected within a few feet of one another, one was compact and maroon and the other was from a powdery gray lens, possibly weathered volcanic ash. The first contained approximately 4 times as many magnetic spherules as the second, probably reflecting a difference in the rates at which the deposits were laid down. Another interesting feature is the high count in the Upper Mancos shale 10 feet below the top of the formation. After the first high count (156 spherules/g) was obtained, a second sample from the same hand specimen was processed. The second sample gave a count (128 spherules/g) reasonably consistent with the first.

The variations in spherule content in the widely spaced samples of the Mancos shales suggest that it might be interesting to make a detailed study of the variations in a limited column, and to correlate them with the pronounced variations in texture seen in these shales. It is hoped to make this the subject of a future report.

Size distributions for the magnetic spherules are of interest; by totaling the counts for the 10 samples of the Mancos shales, sufficient data exist for a frequency curve, shown as the solid line in Figure 1, plotted for diameter intervals of  $5\ \mu$ . For comparison, the dotted frequency curve gives a typical size distribution for the current rate of deposit of black, magnetic spherules from the atmosphere. The frequency scales have been translated to give superposition for

the smaller sizes. The data for the dotted curve were obtained from the deposit in a tray, containing a little glycol, exposed at an isolated point in west-central New Mexico. The deposit over a 7-month period, December to July 1956, is used. The same method of magnetic separation used on the sedimentary material was employed to separate the magnetic fraction of the atmospheric deposit from general dust. The total number of spherules represented by the solid curve is 1320; by the dotted curve 16,900. As far as the present data go, the distributions are practically identical; the

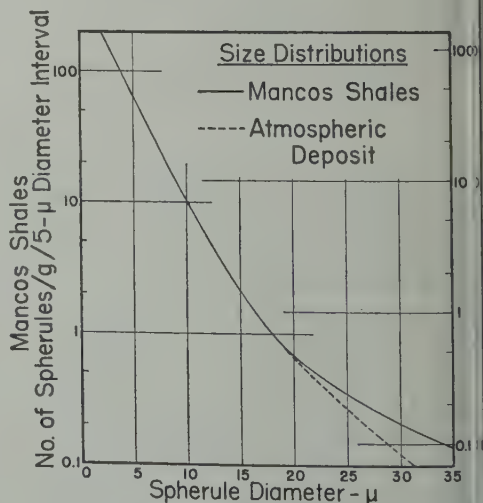


Fig. 1. Size distributions of black, magnetic spherules.

TABLE 3. Content of Black, Magnetic Spherules in Various Clays

| Origin of Clay | Age           | Sample Size, g | Number of Magnetic Spherules per Gram |                   |                    |                    |           | Total |
|----------------|---------------|----------------|---------------------------------------|-------------------|--------------------|--------------------|-----------|-------|
|                |               |                | <5 $\mu$                              | 5 $\mu$ -10 $\mu$ | 10 $\mu$ -15 $\mu$ | 15 $\mu$ -20 $\mu$ | >20 $\mu$ |       |
| Agustin No. 1  | ca. 1000 yr   | 4.1            | 100                                   | 49                | 8                  | 1.2                | 2.2       | 160   |
| Agustin No. 2  | ca. 1000 yr   | 2.0            | 67                                    | 34                | 6                  | 2.0                | 1.5       | 110   |
| Chicago No. 1  | ca. 12,000 yr | 1.9            | 22                                    | 7                 | 3                  | 1.6                | 0.5       | 34    |
| Chicago No. 2  | ca. 12,000 yr | 2.1            | 34                                    | 17                | 1                  | 0.5                | ...       | 53    |
| Popotosa       | Tertiary      | 1.1            | 78                                    | 4                 | ...                | ...                | ...       | 82    |
| Deep sea*      | Recent        | 1.5            | 92                                    | 53                | 8                  | 1.3                | 4.7       | 159   |
| Deep sea†      | Recent        | 6.1            | 103                                   | 54                | 8                  | 1.5                | 1.3       | 168   |

\* From 42°42.6'N, 142°13'W, at 4300 meters depth.

† From 44°31'S, 127°14'W, at 4675 meters depth.

repancy in the 20- to 35- $\mu$  diameter interval not very significant because the solid curve in this interval is based upon only 22 spherules.

*Clays.* In Table 3 are shown the results of magnetic spherule counts on various clays, all of recent age except one from the Popotosa formation which is of Tertiary age.

The study of the San Agustin (N. Mex.) clay originally was undertaken in an attempt to collect magnetic spherules free from the suspicion of being of industrial origin, which suspicion frequently arises in connection with the current spherule deposit from the atmosphere. The San Agustin clay is a playa deposit; the samples used for the counts of Table 3 were taken 42 inches below the surface. The size distribution of spherules in the San Agustin clay is the same, within the experimental limitations, as that found for the Mancos shales and for the present deposit from the atmosphere.

According to radiocarbon dating reported by *Elisby and Sears* [1956], the average rate of deposition over the upper 19 feet of the clay is 0.012 inches/year. If we accept this figure as applying at the 42-inch depth, take a weighted average spherule content of 54 spherules/g for spherules larger than 5  $\mu$  in diameter, and take a density of 2.0 g/cm<sup>3</sup> for the dry sediment, we calculate an average numerical rate of infall of 9 spherules/m<sup>2</sup>/hour. For comparison, from the data for the rate of infall shown in Figure 1, for spherules larger than 5  $\mu$  in diameter, we have a rate of infall of approximately 31 spherules/m<sup>2</sup>/hour. The agreement, even in order of magnitude, between these two figures is regarded, in

view of the several grave uncertainties involved, as fortuitous. Some of these uncertainties are (1) the role of reworked material in both the playa deposit and the atmospheric deposit, (2) the uncertainty in rate of sedimentation at the particular horizon sampled (for instance, a suspicion of varves seen in the soft, distorted clay as collected suggested that the rate of sedimentation may have been several times as large as the rate of 0.012 inches/year used in the calculation), (3) a rather high degree of variability in the atmospheric deposits, and (4) the possibility of penetration by late surface deposits, through cracks in the playa surface. (The probability of such penetration to the 42-inch depth of sediments deposited in the last 100 years is considered rather small; the cracks seen in several visits in dry weather were superficial and the clay below the 12-inch depth was very moist.)

In view of the various uncertainties, a precise age for the San Agustin specimens cannot be assigned. It seems reasonable, however, that their age is in the range 700-1000 years to 3500 years. Except for the unknown factor of penetration, it would seem in any case that the specimens were well removed from the possibility of industrial contamination. The results with the Mancos shales, of course, offer a greater certainty with respect to this.

The Lake Chicago (Wisconsin) clay, furnished through the courtesy of Mr. J. D. Buddhue, was of an age established by radiocarbon dating to be about 12,000 years. According to Mr. Buddhue, the varves at this horizon indicated that

this clay may have been laid down at the comparatively rapid rate of 1.25 inch/year.

The spherule count in the Lake Chicago clay appears to be significantly smaller than the count in the San Agustín clay, though the ratio of the counts is not as low as the inverse of the ratio of the apparent rates of deposition. This is not surprising, of course, since all such deposits must contain reworked material which probably contains spherules in addition to those deposited directly from the atmosphere.

Table 3 also shows the counts on two samples of deep sea clay, obtained through the courtesy of Mr. R. W. Rex, Dr. E. D. Goldberg, and Mr. Victor Vacquier, all of the Scripps Institution of Oceanography. Both samples were taken from near the tops of cores from the ocean bottom and are essentially surface samples. Attempts were made to study several other samples, but the presence of a large amount of what appeared to be finely divided volcanic glass, containing dispersed magnetic particles which carried it through the magnetic separation, made the search for the spherules almost impossible. The counts shown in the table are the summations for two lots of sample 1 and for three lots of sample 2.

The size distribution for the magnetic spherules counted in deep sea clays 1 and 2 is shown by the solid-line frequency curve in Figure 2. For spherules of diameters less than  $20\ \mu$ , the size distribution is practically identical with the

size distribution in the Mancos shales (Fig. 1), though the magnitudes of the numbers for the deep sea clays are about 1.7 times as large as the Mancos shales.

The count for all spherule sizes, for a total of 7.6 grams of the deep sea clay, is about 140 spherules/g. This is about 140 times as large as the average of *Laevastu and Mellis* [1955], a discrepancy that obviously calls for discussion. The size distribution obtained by *Laevastu and Mellis*, recalculated from their cumulative percentage curve, is shown as the dotted curve in Figure 2. It is seen that, though the difference between the two curves is very large for the small spherules, they tend to converge as the size of the spherules increases; the two curves blend very well if the upper curve is translated downward by a factor of 2.5, as shown by the dotted curve. A factor of 2.5, of course, can easily be accounted for by variation between samples; for example, the counts in sample 72 of *Laevastu and Mellis* are shown as varying by a factor of 2.5.

It is tentatively suggested that the magnetic separation used by *Laevastu and Mellis* may have discriminated against the smaller spherules. The separation was apparently carried out without previous dispersion and removal of the bulk of the clay, and the manner in which a 'strong electromagnet' was used to collect the magnetic particles from the washings of their Fram screen was not specified.

It is further suggested that the true size distribution for magnetic spherules in the deep sea clay may have a form similar to the dotted curve of Figure 2 for spherules smaller than  $20\ \mu$  in diameter, whereas the dashed curve, representing the data of *Laevastu and Mellis*, gives the distribution for spherules larger than  $400\ \mu$  in diameter.

Calculations of the total mass of the spherules per gram of sediment are not affected appreciably by the use of the combination curve instead of the original curve of *Laevastu and Mellis*. According to *Laevastu and Mellis*, their data indicate an annual infall of about 125 metric tons of the spherules over the surface of the earth. According to *Pettersson and Fredriksson* [1958], the annual infall is 2400 to 5000 metric tons. Whichever estimate is accepted, it seems that the annual infall calculated from the des-

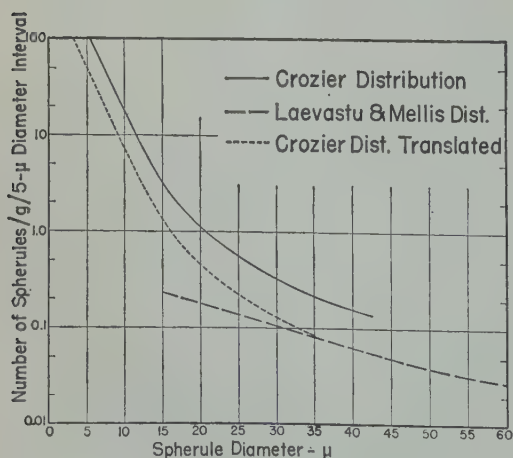


Fig. 2. Size distributions of black, magnetic spherules in deep sea sediments.



sea deposits is small compared with the infall calculated from the playa deposit, which appears to be in the range of 30,000 to 150,000 tons per year, depending on the rate of sedimentation assumed and disregarding the mass contribution from spherules larger than  $40\ \mu$  in diameter. If further work on the playa deposits should show that a large part of the mass is associated with spherules larger than  $40\ \mu$  in diameter, as is the case with the deep sea spherules, the annual infall may be found to be much larger than 150,000 tons.

If the current mass rate of deposition from the atmosphere is calculated on the basis of such data as those in Figure 1, a figure of the order of 150,000 tons per year is obtained. Hodge and Wildt [1958] calculated an annual infall of 500,000 tons, using only spherules of diameters from  $3\ \mu$  to  $15\ \mu$ .

It is not proper, of course, to draw elaborate conclusions from such a small sampling, but the above results would seem to raise a question concerning the survival of magnetic spherules in recognizable form in the deep sea. In this connection it is interesting to note that iron-nickel, often as free metal, seems to be prominent in analyses of the deep sea spherules, for example, the analyses of Smales, Mapper, and Wood [1958] and of Castaing and Fredriksson [1958]. On the other hand, a preliminary investigation by Crozier and Chen [1957] indicated that nickel is quite rare in spherules collected from the atmosphere. The possibility is then apparent that a small minority of iron-nickel spherules in the atmospheric infall represents the principal part of the spherule content of the deep sea sediments. Perhaps the plain magnetite spherules do not survive well in the sea.

### CONCLUSIONS

This work appears to demonstrate that black, magnetic spherules, of probable meteoritic origin, are widely present in sedimentary material. Large discrepancies exist between the apparent rates of deposition of the spherules in the deep sea and in the San Agustin playa deposit, and further work on the quantities, as well as on the physical and chemical nature of the spherules, is needed to throw light on these discrepancies. Ultimately, perhaps, the spherule content may

be a useful tool in the study of rates of sedimentation. Work with larger samples of sediment needs to be done to establish size distributions and mass contributions of spherules in the larger size ranges.

*Acknowledgment.* This work was supported by the Office of Naval Research under Contract Nonr-815(02).

### REFERENCES

- Castaing, R., and K. Fredriksson, Analyses of cosmic spherules with an x-ray microanalyser, *Geochim. et Cosmochim. Acta*, **14**, 114-117, 1958.
- Clisby, K. H., and P. B. Sears, San Agustin Plains—Pleistocene climatic changes, *Science*, **124**, 537-538, 1956.
- Crozier, W. D., Recent developments in the collection and identification of meteoritic dust, Paper presented at the Los Angeles meeting of The Meteoritical Society, September 4, 1957.
- Crozier, W. D., Black, magnetic spherules in sedimentary rocks and unconsolidated clays, Paper presented at the Winslow, Ariz., meeting of The Meteoritical Society, August 31, 1958.
- Crozier, W. D., and Kung-Yung Chen, An attempt to detect nickel in airborne particles of presumed meteoritic origin, Paper presented at the Tucson, Ariz., meeting of the Southwestern and Rocky Mountain Division of the Am. Assoc. Advance. Sci., May 1, 1957.
- Ersine, W. S., "Micrometeorites" of the Todilto gypsum—a preliminary investigation, Paper presented at the Albuquerque, N. Mex., meeting of the N. Mex. Geol. Soc., April 24, 1959.
- Hodge, P. W., and R. Wildt, A search for airborne particles of meteoritic origin, *Geochim. et Cosmochim. Acta*, **14**, 126-133, 1958.
- Laevastu, T., and O. Mellis, Extraterrestrial material in deep-sea deposits, *Trans. Am. Geophys. Union*, **36**, 385-389, 1955.
- Murray, Sir John, On the distribution of volcanic debris over the floor of the ocean—its character, source, and some of the products of its disintegration and decomposition, *Proc. Roy. Soc. Edinburgh*, **9**, 258, 1876.
- Pettersson, H., and K. Fredriksson, Magnetic spherules in deep-sea deposits, *Pacific Sci.*, **12**, 71-81, 1958.
- Skolnick, H., Miocene meteoritic dust, Paper presented at the Missoula, Mont., meeting of the Geol. Soc. Am., Rocky Mountain Section, May 15, 1959.
- Smales, A. A., D. Mapper, and A. J. Wood, Radioactivation analysis of "cosmic" and other magnetic spherules, *Geochim. et Cosmochim. Acta*, **13**, 123-126, 1958.

(Manuscript received April 21, 1960; revised May 27, 1960.)





## Principles of Structural Stability of Underground Salt Cavities

SHOSEI SERATA

*Department of Civil Engineering  
Michigan State University  
East Lansing, Michigan*

AND

EARNEST F. GLOYNA

*Department of Civil Engineering  
The University of Texas  
Austin, Texas*

**Abstract.** The principles of structural stability of underground salt cavities and the significance of the principles as they relate to other cavities are discussed. The theory of plasticity is applied to the evaluation of stress and strain conditions of the salt cavities. The concept of a yielded zone which develops around the cavities is introduced, and a theoretical development of the extent and stress distribution of the zone is illustrated through the use of ideal spherical and cylindrical cavities under uniform triaxial compression. Applicability of the concept to actual conditions, such as cavity irregularities, brittleness of formation, and nonhydrostatic loading, is discussed.

**Introduction.** As has often been observed, underground cavities generally do not collapse because of direct earth pressure. This ability of the cavities in general to resist failure cannot be satisfactorily explained by the common concept of elasticity as applied to ground formations. A fundamental understanding of the structural behavior of underground salt cavities is now needed because there is increased interest in the long-term use of the cavities. In addition to using the underground cavities as storage containers for liquefied petroleum products and other materials, there is also the possibility that certain cavities may be used for storing concentrated radioactive wastes, possibly for hundreds of years. In studies of underground structures, the common concepts of the strength of materials and experimental data on ground formations obtained by ordinary testing methods are not applicable for the following reasons:

1. Underground formations are always subject to three-dimensional (triaxial) compression; therefore the actual strength of a ground formation should be defined in three-dimensional terms.

2. Under triaxial compression, the structural behavior of a material differs from that observed in ordinary, two-dimensional testing.

3. The plastic and viscous nature are of a material particularly pronounced in triaxial compression; therefore the strength of a material should be expressed in terms of time.

4. Internal friction in a triaxially compressed material, in contrast to ordinary compression, significantly alters its structural behavior.

The purpose of this study is to establish theoretical principles governing the strength and stability of underground salt cavities. Plasticity and viscosity of the formation are considered in addition to elasticity. Finally, the significance of the principles of salt cavity deformation to cavities in other formations is discussed.

*Limitations to the theory of elasticity.* The creation of a cavity in a stressed elastic medium results in the redistribution of the initial stress because of the concentration of greater shearing stress around the cavity. The magnitude of the stress concentration can be determined precisely

by either theoretical or experimental methods for any form of single or multiple cavities under any external pressure conditions.

There would be no structural problems if underground formations behaved like elastic or even near-elastic bodies. In reality, however, all geological formations exhibit both plastic and viscous properties, depending upon overburden pressures and the triaxial properties of the formation. Thus, the theory of elasticity can be applied only to cavities of certain formations of limited depths. Calculations based on the theory of elasticity show that the maximum shearing stress around cylindrical and spherical cavities always appears on the boundary. When this maximum shearing stress exceeds the maximum shearing strength of the formation, the material on the boundary should fail, according to theory.

In many formations this failure does not occur, because the cavity boundary would yield before the shearing stress reached the maximum strength. A release of stress on the boundary and a transfer of the concentrated stress to the formation adjacent to the boundary would result, as is illustrated in Figures 5 and 6. This yielding characteristic of rock salt is particularly pronounced when the material is in triaxial compression. In order to evaluate this deformation, the triaxial properties of the material should be known.

*Triaxial behavior of brittle material.* Rock salt exhibits brittle failure under ordinary compression. Under triaxial compression, however, the slightest yielding of the brittle material along the plane of the maximum shearing stress causes a frictional resistance force in the same plane. This frictional resistance seems to be the mechanism by which a brittle ground formation like rock salt behaves plastically under triaxial compression. Since the friction force is directly proportional to the normal stress acting perpendicularly to the plane of the maximum shearing stress, the plastic behavior of the brittle material is primarily dependent upon the magnitude and geometry of the triaxial compression. The degree of triaxial compression may be indicated by the mean principal stress  $\sigma_m$ , which is an average of the three principal stresses  $\sigma_1, \sigma_2, \sigma_3$ .

Therefore

$$\sigma_m = \frac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3) \tag{1}$$

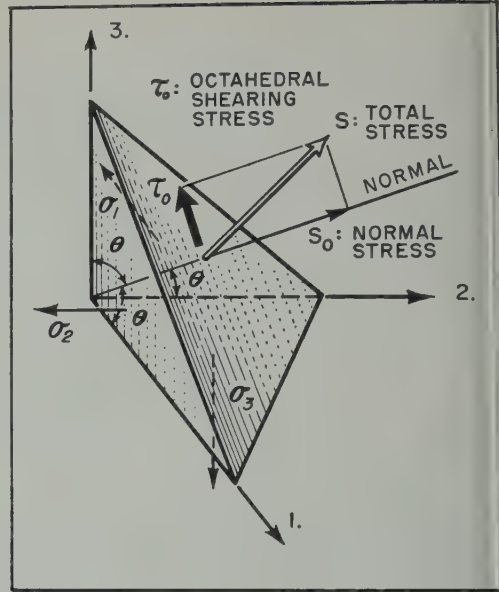


Fig. 1. Octahedral shearing stress in principal stress coordinate.

For rock salt, a  $\sigma_m$  value in excess of 5000 psi is sufficient to produce the necessary friction force to prevent brittle fractures.

Under triaxial pressure the strength of rock salt is dependent upon the maximum triaxial shearing stress occurring in the material. The maximum shearing stress is called octahedral shearing stress,  $\tau_o$ , which can be expressed as

$$\tau_o = \frac{1}{3} \sqrt{(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2} \tag{2}$$

The octahedral shearing stress is graphically presented in Figure 1.

For any given material, the octahedral shear

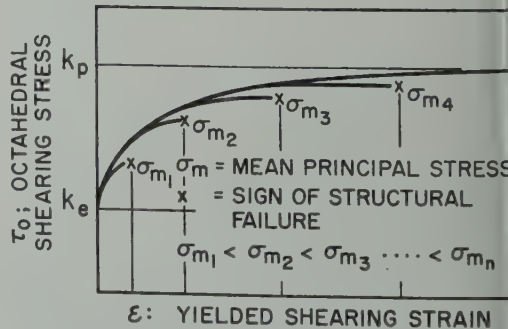


Fig. 2. Increasing plasticity of brittle material with increase of triaxial compression.

strength can be determined by a laboratory triaxial test, if sufficient triaxial pressure is provided. The general relation between the octahedral shearing stress  $\tau_o$  and the yielding shear strain  $\epsilon$  is shown in Figure 2.

As is seen in Figure 2, rock salt starts to yield whenever  $\tau_o$  exceeds the octahedral elastic limit of the material. A further increase of  $\tau_o$  above the  $K_e$  point accelerates the strain, as Figure 2 illustrates. Consequently, the octahedral shearing stress cannot exceed a certain stress value defined as the plastic limit  $K_p$  of the material. Beyond this limit, the material deforms infinitely. On the basis of this rather unusual fact, the yielding deformation of rock salt can be treated as plastic deformation.

Ground formations may not have definite values of  $K_p$ , as the triaxial pressure requirements needed to exceed their plastic limits are usually not obtainable in the laboratory. However, their plasticity increases with increase of the triaxial compression.

**Creep flow.** In addition to instantaneous yielding deformation, rock salt is also subject to pseudoviscous flow. Pseudoviscous flow is defined as gradual creep deformation when constant pressure is applied below the elastic limit. This viscous deformation in salt can be accumulated in considerable quantity for a long period of constant loading. In principle, the octahedral yielding strength  $K_y$  decreases with increase of the observation period due to this effect. Generally, the yielding strength  $K_y$  has the characteristic of being a logarithmic function of time. Because of this viscous nature of formations, a yielded zone continues to extend with decreasing  $K_y$ . In the design of an underground cavity, the  $K_y$  value should be determined as a function of time in order that the extent of the expected yielded zone for a given period of time may be calculated.

**Initially existing ground pressure.** The initially existing ground pressure can be determined by using Poisson's ratio of the formation for relatively shallow depths where the formation is nearly elastic, provided no tectonic pressure exists. Beyond this depth the ground pressure becomes independent of Poisson's ratio but dependent on the cavity depth and the octahedral yielding stress  $K_y$ . For most ground formations the lateral earth pressure is more than one-third

of the overburden pressure, if Poisson's ratio is assumed to be greater than 0.25. However, the ratio of the lateral to the overburden pressure approaches unity with further increase of the depth. Should tectonic pressure be present, the lateral pressure can be either greater or less than the vertical. However, the possible range of the lateral pressure remains within certain limits. This range is rather restricted in rock salt owing to its small  $K_p$  value.

**Development of yielded zone around spherical cavity.** As is shown in the previous discussion, a yielded zone may develop around an underground salt cavity regardless of strength and brittleness of the formation. For simple forms of cavities, such as spheres or cylinders, the development of the zone can be mathematically determined and the structural conditions can be mathematically formulated. Although no actual cavity will be ideal, the theory derived from an ideal condition discloses fundamental principles regarding cavity structures. The principles obtained from an ideal condition can be adapted for the evaluation of nonideal conditions. A spherical cavity subject to uniform external pressure is considered here as an example.

A yielded zone starts to develop around the spherical cavity when the octahedral shearing stress  $\tau_o$  exceeds the elastic octahedral shearing strength  $K_e$ . The stress redistribution in this yielded zone can be obtained by introducing the yielding octahedral shearing strength  $K_y$  into the differential equation of stress equilibrium in the zone (Fig. 3). Therefore:

$$d\sigma_r/dr + (2/r)(\sigma_r - \sigma_t) = 0 \quad (3)$$

The octahedral shearing stress around the cavity is obtained by substituting  $\sigma_1 = \sigma_r$  and  $\sigma_2 = \sigma_3 = \sigma_t$  into equation 2 as

$$\tau_o = \pm(\sqrt{2}/3)(\sigma_r - \sigma_t) \quad (4)$$

Thus, (3) can be rewritten for the stress condition in the yielded zone:

$$\begin{aligned} \frac{d\sigma_r}{dr} &= -\frac{2}{r}(\sigma_r - \sigma_t) = \pm 3\sqrt{2} \frac{|\tau_{0\max}|}{r} \\ &= \pm 3\sqrt{2} \frac{K_y}{r} \end{aligned}$$

At the same time, the stress condition around



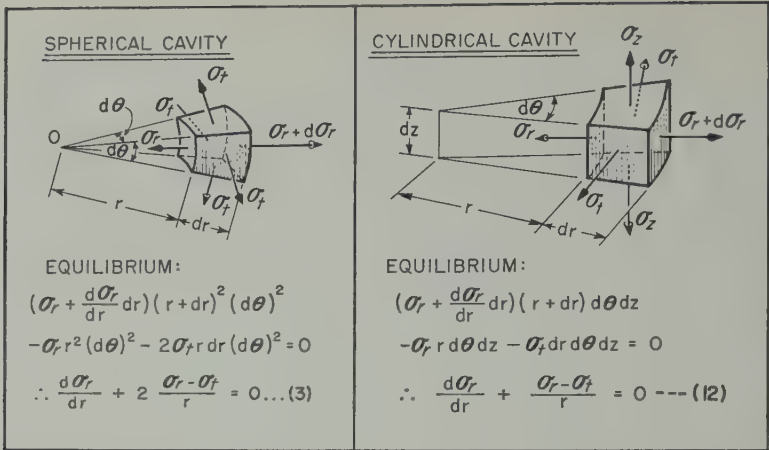


Fig. 3. Dimensions and stress equilibrium of volume elements around spherical and cylindrical cavity

the cavity is restricted to

$$\sigma_t < 0, \quad \sigma_r < 0, \quad \sigma_t < \sigma_r,$$

and  $d\sigma_r/dr < 0$

Therefore, the fundamental stress equations of the yielded zone are expressed as

$$d\sigma_r/dr = -3\sqrt{2} K_y/r \quad (5)$$

and

$$\sigma_r - \sigma_t = (3/\sqrt{2}) K_y \quad (6)$$

By integration the general equations of stress distribution in the yielded zone are obtained:

$$\sigma_r = C - 3\sqrt{2} K_y \ln r \quad (7)$$

$$\sigma_t = C - 3\sqrt{2} K_y (\ln r + 0.5)$$

where  $C$  is a coefficient of integration.

Surrounding this yielded zone is a zone of elastic state extending to an infinite distance in the formation. The general equation of stress distribution in this elastic zone is obtained from the theory of elasticity:

$$S_r = A - B/r^3 \quad (8)$$

$$S_t = A + B/2r$$

The coefficients of integration,  $A$ ,  $B$ , and  $C$  in (7) and (8), can be determined simultaneously from the boundary conditions of the cavity by application of the following boundary conditions:

At the wall of the cavity  $r = a$ ,  $\therefore \sigma_r = -P_i$

At an infinite distance away from the cavity  $r = \infty$ ,  $\therefore S_r = -P_o$

At the borderline between yielded and elastic zones  $r = \rho$ ,  $\therefore S_r = \sigma_r$

The integration constants can thus be obtained, and the following stress distribution around the cavity is derived:

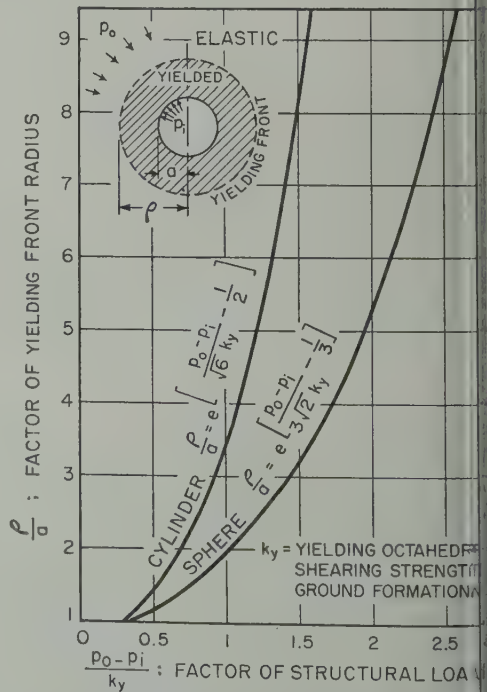


Fig. 4. Development of yielded zone around spherical and cylindrical underground cavity

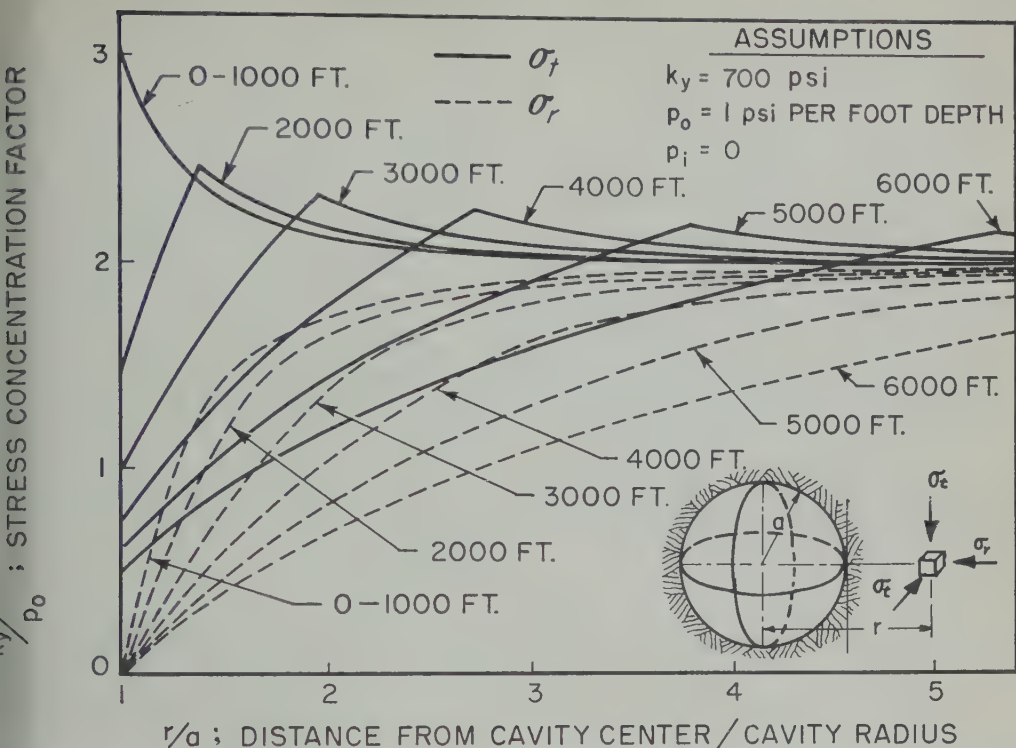


Fig. 5. Stress distribution in yielded and elastic zones around spherical underground cavity at various depths.

the yielded zone ( $a \leq r \leq \rho$ ),

$$\sigma_r = -P_i - 3\sqrt{2} K_v \ln r/a \quad (9)$$

$$\sigma_i = -P_i - 3\sqrt{2} K_v (\ln r/a + 0.5)$$

where  $P_i$ ,  $P_o$  = internal and external pressures  
 the cavity, respectively.

the elastic zone,

$$\sigma_r = -P_o$$

$$+ \frac{\rho^3}{r^3} (P_o - P_i - 3\sqrt{2} K_v \ln \rho/a) \quad (10)$$

$$\sigma_r = -P_o$$

$$+ \frac{\rho^3}{2r^3} (P_o - P_i - 3\sqrt{2} K_v \ln \rho/a)$$

The radius to the front of the yielded zone,  $\rho$ ,  
 can be obtained as a function of the loading  
 condition to the cavity. Since  $\sigma_r = S_r$  and  
 $\sigma_i = S_i$  at  $r = \rho$ , equation 6 can be rewritten  
 follows at  $r = \rho$ :

$$S_i - S_r = -\frac{3}{\sqrt{2}} K_v$$

By substitution of (10) into the above relation,  
 the radius of the yielded front is obtained as:

$$\rho = a \exp \left( \frac{P_o - P_i}{3\sqrt{2} K_v} - \frac{1}{3} \right) \quad (11)$$

In Figure 4, the radius of the yielded front  
 is given under various stress conditions of the  
 cavity. The stress distribution in both the  
 yielded and the elastic zones can be calculated  
 by substituting the front radius of the yielded  
 zone  $\rho$  into (9) and (10). As a sample calculation,  
 the stress distribution around the cavity is  
 graphically presented in Figure 5 for a formation  
 in which  $K_v = 700$  psi and  $P_o = 1$  psi per foot  
 depth.

*Development of yielded zone around cylindrical  
 cavity.* Development of the yielded zone and  
 the stress distribution around a cylindrical salt  
 cavity was analyzed by a method similar to

that used for a spherical cavity. The differential equation of stress equilibrium in a thick-walled cylinder (Fig. 3) is

$$d\sigma_r/dr + (\sigma_r - \sigma_t)/r = 0 \quad (12)$$

The octahedral shearing stress around the cylindrical cavity, from (2), is

$$\tau_0 = \pm \frac{1}{3} \sqrt{(\sigma_r - \sigma_t)^2 + (\sigma_t - \sigma_z)^2 + (\sigma_z - \sigma_r)^2}$$

Since a displacement in the axial direction in the yielded zone is not probable, the axial yielded strain can be equated to zero. The axial yielded strain  $\epsilon_z$  can be expressed as

$$\epsilon_z = F(\tau_0)[\sigma_z - \frac{1}{2}(\sigma_t + \sigma_r)]$$

where  $F(\tau_0)$  is a characteristic function of  $\tau_0$  and is not zero. Therefore,  $\epsilon_z = 0$  leads to

$$\sigma_z = (\sigma_t + \sigma_r)/2$$

By substitution of  $\sigma_z$  into the above equation of octahedral shearing stress the stress in a yielded zone is obtained:

$$\tau_0 = \pm(1/\sqrt{6})(\sigma_r - \sigma_t) \quad (13)$$

Thus (12) can be rewritten for the stress condition in the yielded zone:

$$d\sigma_r/dr = (-1/r)(\sigma_r - \sigma_t) \\ = \pm(\sqrt{6}/r) |\tau_{0\max}| = \pm(\sqrt{6}/r) K_y$$

At the same time, the stress condition around the cavity is restricted to

$$\sigma_t < 0, \quad \sigma_r < 0, \quad \sigma_t < \sigma_r,$$

$$\text{and } d\sigma_r/dr < 0$$

Therefore, the fundamental stress equations of the yielded zone around a cylindrical cavity are expressed as

$$d\sigma_r/dr = -\sqrt{6} K_y/r \\ \sigma_r - \sigma_t = \sqrt{6} K_y \quad (14) \\ \sigma_z = \frac{1}{2}(\sigma_r + \sigma_t)$$

These three equations can be simultaneously solved by using the method similar to that for the spherical cavity. Thus, the equations for the stress distribution in both the yielded and the elastic zones around the cavity are found:

In the yielded zone ( $a \leq r \leq \rho$ ),

$$\sigma_r = -P_i - \sqrt{6} K_y \ln r/a \\ \sigma_t = -P_i - \sqrt{6} K_y (\ln r/a + 1) \\ \sigma_z = -P_i - \sqrt{6} K_y (\ln r/a + 0.5) \\ \text{In the elastic zone } (\rho \leq r \leq \infty),$$

$$S_r = -P_0 \\ + (\rho^2/r^2)(P_0 - P_i - \sqrt{6} K_y \ln \rho/a) \\ S_t = -P_0 \\ - (\rho^2/r^2)(P_0 - P_i - \sqrt{6} K_y \ln \rho/a) \\ S_z = -P_0$$

The radius to the front of the yielded zone derived as a function of the stress condition around the cavity. Since  $\sigma_r = S_r$  and  $\sigma_t = S_t$  at  $r = \rho$ , equation 14 becomes

$$S_r - S_t = \sqrt{6} K_y$$

Substituting  $S_r$  and  $S_t$  of (16) into the above equation, we find the relation between the radius of the yielded front and the loading condition:

$$\rho = a \exp \left( \frac{P_0 - P_i}{\sqrt{6} K_y} - \frac{1}{2} \right)$$

In Figure 4, the front radius of the yielded zone under various loading conditions of the cylindrical cavity is given as a function of the internal and external pressures and the yielding octahedral shearing stress.

The stress distribution in the yielded and elastic zones around the cylindrical cavity is calculated by substituting the front radius into (15) and (16). The stress distribution around the cavity is illustrated in Figure 6 assuming  $K_y = 700$  psi and  $P_0 = 1$  psi per inch depth.

*Field application of the theory.* The theory developed here by assuming ideal conditions can be successfully applied to actual field conditions. The application, of course, requires careful adaptation of the theory to field conditions. First, there are three important factors which require consideration: shape of cavity, in-situ stress conditions, and the yielding octahedral shearing strength of the formation.

Unlike cavities in an elastic medium, the shape is not too significant for the structural stability of an underground salt cavity. In any irregular

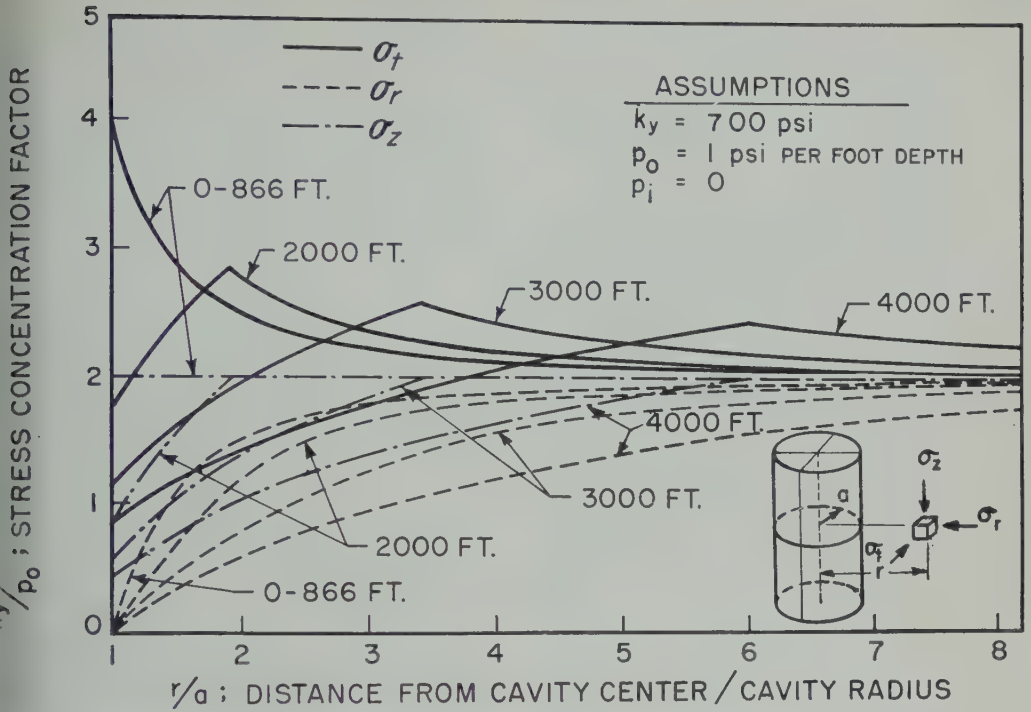


Fig. 6. Stress distribution in yielded and elastic zones around cylindrical underground cavity at various depths.

...vity the yielding first appears at sharp corners, where higher elastic stresses are concentrated. With an increase of the external pressure the yielded zones expand around the corners, where the distribution of the zone is quite different from that of ideal cavities. However, the distribution around irregular cavities becomes similar to that of the ideal ones with further development of the zone.

When the external pressure is not hydrostatic the development of the yielded zone is not uniform around the opening. The distribution of the yielded zone is affected by the ratio of the vertical and lateral pressures. The yielded zone appears first around walls which are subject to the minimum principal pressure. With an increase of the pressure, however, the yielded zone develops all around the cavity. With an extended yielded zone, the distribution becomes similar to that of the cavity with uniform pressure. Because of this yielded zone, the nonhydrostatic pressure has little effect upon the structural stability of the cavity. The ulti-

mate form of the yielded zone approaches an ellipse or ellipsoid whose ratio of major to minor axis is equal to the ratio of major to minor principal stresses. One local area of weakness of the cavity under nonideal conditions occurs on the boundary between the zones of initial yielding. As the yielded zones surround those areas, they become isolated from the external pressure. This process causes roof failure and wall and floor fracture.

The most important structural property of rock salt is its yielding octahedral shearing strength  $K_y$ . Analysis of reported triaxial tests of rock salt by some investigators indicates a rather wide variation of  $K_y$  value ranging from 1000 to 4000 psi at room temperature. It appears that further study of the triaxial strength properties of rock salt is needed. Furthermore, the value of  $K_y$  should be determined as a function of time because rock salt is not ideally plastic and is therefore subject to considerable pseudo-viscous flow.

Reduction of the cavity space is anticipated



when a yielded zone develops around the cavity because of a certain volume expansion in the yielded mass. This reduction can be calculated from the mass in the yielded zone and its coefficient of expansion. This coefficient can be obtained by laboratory tests as a function of  $K_y$  and the initial stress.

The principles of salt cavities outlined in this paper have been verified qualitatively both in the laboratory and in underground salt mines having different types of rock salt formations. Analyses of all the experimental results have provided strong support for the theoretical conclusions presented here. The details of this experimental verification will be discussed in a separate article.

*Significance of the principles to cavities of other formations.* Application of the principles developed for salt cavities to cavities of other underground formations is restricted because of a unique property of rock salt: it becomes nearly plastic with rather low  $\sigma_m$  and  $K_y$  values. Most other ground formations seldom reach a plastic condition under ordinary laboratory triaxial testing pressures. The conditions described by the curves of  $\sigma_m$  to  $\sigma_{m_s}$  in Figure 2 would hold true for some rocks even under laboratory pressures of over  $10^6$  psi. In these cases, naturally, the theory of elasticity is more reasonably applied than the principles developed here.

However, it is possible that with an increase of  $\sigma_m$  an underground material would eventually approach the plastic condition characterized by the horizontal portion of the envelope curve of Figure 2, indicating a certain  $K_y$  value. Evidence of extremely curved rock formations may support these principles.

*Conclusions.* From the theoretical study discussed herein, the following conclusions may be drawn:

1. Cavities created in underground salt formations may shrink but do not collapse under direct earth pressure regardless of their depth because of a yielded zone which develops around the cavities.

2. Around the salt cavities, a yielded zone develops due to an increase in plasticity resulting from an increase in triaxial pressure.

3. An underground salt cavity establishes its structural equilibrium by expanding its yielded

zone sufficiently to withstand the external earth pressure.

4. For salt cavities of simple form, such as a sphere or a cylinder, the front radius of the yielded zone,  $\rho$ , can be calculated as a function of the yielding octahedral shearing strength and external and internal pressures  $P_o$ ,  $P_i$ , and initial radius of the cavity.

For a spherical cavity,

$$\rho = a \exp \left( \frac{P_o - P_i}{3\sqrt{2} K_y} - \frac{1}{3} \right)$$

For a cylindrical cavity

$$\rho = a \exp \left( \frac{P_o - P_i}{\sqrt{6} K_y} - \frac{1}{2} \right)$$

5. The most important structural property of rock salt formation is its yielding octahedral shearing strength  $K_y$ , which can be determined in a laboratory.

6. The amount of cavity volume reduction can also be predicted if the expansion coefficient of the ground formation is determined as a function of triaxial pressure.

7. A salt cavity of irregular form under nonhydrostatic compression would probably be as stable as an ideal cavity owing to the pronounced yielding deformation in triaxial compression. The ultimate structural equilibrium of an irregularly shaped cavity can be approximated by comparisons with ideal spherical and cylindrical cavities.

8. The loading condition of an underground salt cavity is best described by the term 'factor of structural loading,  $x$ ,' which is defined as

$$x = (P_o - P_i)/K_y$$

where  $P_o$  = overburden pressure of the cavity

$P_i$  = internal pressure of the cavity

$K_y$  = yielding octahedral shearing strength of the formation.

9. Elastic theory can be applied for cavity stress analysis only when the maximum octahedral shearing stress remains less than the elastic octahedral shearing strength,  $K_e$ . Otherwise, the theory of plasticity must be introduced.

10. The principles developed for salt cavities may be applicable to cavities of some underground formations other than salt if the existing

mean triaxial pressure on the cavities exceeds the minimum mean triaxial pressure required for plastic behavior of the formations.

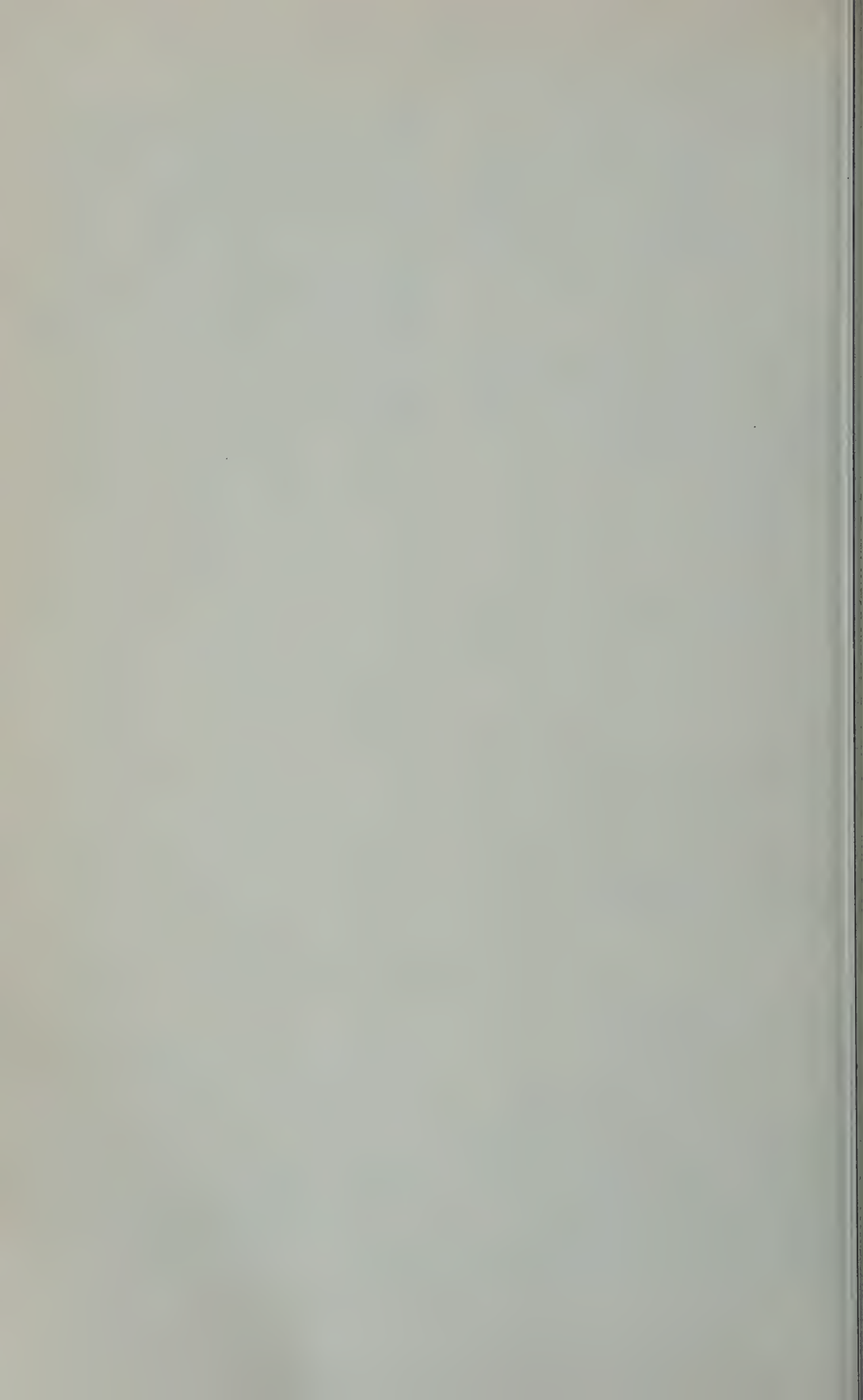
11. For underground formations which do not exhibit sufficient plasticity for existing triaxial compression, the structural stability of their cavities may be reasonably estimated as an intermediate condition between the two calculable extreme conditions, namely, ideally elastic and ideally plastic.

*Acknowledgment.* This study was sponsored by the U. S. Atomic Energy Commission, under the guidance of Dr. Joseph A. Lieberman, Sanitary Engineer, Division of Reactor Development. This effort has been part of the Reactor Fuel Waste Disposal Project currently being investigated at the University of Texas Sanitary Engineering Research Laboratory.

## REFERENCES

- Handin, J., An application of high pressure in geophysics, *Trans. ASME*, 75, 315-324, 1953.  
Heywood, R. B., *Design by Photoelasticity*, 1 and 2, McGraw-Hill Book Co., New York, 1941.  
Hoffman, O., and G. Sachs, *Introduction to the Theory of Plasticity for Engineers*, McGraw-Hill Book Co., New York, 1953.  
Nadai, A., *Theory of Flow and Fracture of Solids*, McGraw-Hill Book Co., New York, 1950.  
Serata, S., and E. F. Gloyna, Development of design principle for disposal of reactor fuel waste into underground salt cavities, *Tech. Rept. of Univ. Texas to Atomic Energy Comm.*, January 1, 1959.  
Timoshenko, S., and J. N. Goodier, *Theory of Elasticity*, McGraw-Hill Book Co., New York, 1951.

(Manuscript received February 20, 1960; revised May 31, 1960.)



## Geomagnetic and Solar Data

J. VIRGINIA LINCOLN

*Central Radio Propagation Laboratory  
National Bureau of Standards  
Boulder, Colorado*

### INTERNATIONAL DATA ON MAGNETIC DISTURBANCES

This report continues the series which has appeared regularly in this JOURNAL since vol. 4 (3), 295, 1949. Please refer to that first report for an explanation of the data given, and to vol. 59 (3), 423, 1954, for the definition of *Ap*.

*Note:* Additional and final 'Geomagnetic and Solar Data' appear in due course in the following international publications: *Quarterly Bulletin on Solar Activity*, International Astronomical Union, c/o Eidgen. Steinwarte, Zurich, Switzerland; *AGA Bulletins, Geomagnetic Indices K and C*, c/o J. Bartels, A. Romaña, and J. Veldkamp, International Union of Geodesy and Geophysics, Association of Geomagnetism and Aeronomy, c/o V. Laursen, Meteorologisk Institut, Charlottenlund, Denmark.

### SUDDEN COMMENCEMENTS AND SOLAR-FLARE EFFECTS, FOURTH QUARTER, 1959

#### Preliminary Report of Sudden Commencements

S.c.'s given by ten or more stations are in italics. Times are mean values obtained from normal magnetograms.

#### *Sudden commencements followed by a magnetic storm or a period of storminess (s.s.c.)*

1959 October 03d 17h 11m: Es VI Ci.—04d 8h 56m: Cm (sfe: MB?).—05d 16h 04m: Mb Ki Ka Ky Hn Bi To (si: Es).—17d 06h 24m: K MB.—29d 23h 47m: fifty-seven (ssc: 56; si: 1).

1959 November 13d 11h 19m: Es Lg Eb Tl IB (pt: Cm CF; sfe: El).—21d 11h 36m: seventeen (ssc: 12; si: 5).—27d 23h 51m: fifty-nine.—30d 05h 53m: Cm Ag IK MB AA? (bs: Es).—30d 06h 53m: Es Wn Pr CF SF (si: So Do Cm Ma).

1959 December 05d 06h 59m: fifty-seven.—13d 18h 22m: Pa Hn Hr? (si: Mc Va; bs: Bn).—23d 15h 25m: forty-nine (ssc: 38; si: 11).—26d 10h 22m: VI Cm.

#### *Sudden impulses found in the magnetograms (s.i.)*

1959 October 02d 10h 40m: Ma Lg IK Qu Ta Mc Bn El Hr.—03d 07h 40m: Pi (ssc: To).—04d 12h 54m: Es (ssc: Le; bs: Wn).—06d 01h 54m: IK Bn El (pg: Bi).—17d 09h 12m: Ba (ssc: Cm).—21d 16h 05m: thirty-two (si: 28; ssc: 1; pt: 2; sfe: Gu).—22d 00h 50m: Gn (bs: Ta).—30d 12h 32m: Fr (ssc: Ci; pt: Cm).—30d 14h 34m: Cm Mb Fr Ka Ky Pi (ssc: Le).

1959 November 02d 18h 29m: Es Db Pr IK (bps: CF Qu).—02d 20h 08m: Es (pt: VI).—02d 23h 25m: Es (pt: CF).—04d 07h 13m: Co (ssc: Es).—06d 11h 42m: MB Hn Bi Va.—09d 03h 18m: Ma CF Su IK Eb Bn Hr Tw; ssc: Cm Lg Md Tl Ta Gu Pa Hn; bs: Mc To.—14d 13h 18m: Ma Db (pg: Cm El).—22d 08h 25m: Qu Ta Mc Lr.—26d 21h 42m: twenty-eight (si: 23; ssc: 1; bps: 1; sfe: PM? Hu).—28d 08h 29m: Wn (pg: Cm).—28d 09h 38m: Cm Lr.—28d 17h 47m: Es Qu Lr.—30d 17h 33m: Es IK Fr.

1959 December 03d 21h 19m: Es (bs: VI).—04d 04h 32m: Mb Ka Ky El To.—05d 11h 36m: Qu (pt: Cm).—05d 13h 10m: Fr MB Bi To (pt: Cm).—05d 18h 47m: Es Ta Ho AA MB Ba Bi (pg: Cm).—05d 20h 49m: IK (pt: Cm).—06d 05h 19m: Ma AA? (pt: Cm).—08d 07h 55m: Ks AA Hr.—13d 21h 48m: VI MB Tn (ssc: Gu Gn; bp: Lg Tl).—13d 22h 09m: Es IK MB Mc El.—24d 06h 24m: Es Ma Ks (sfe: Db).—26d 11h 53m: Es (ssc: MB).—27d 11h 02m: Es Cm (bp: Ho).—28d 11h 50m: Cm (ssc: Le).—28d 13h 26m: Es Fu MB.—30d 06h 09m: Hr (bs: To).

#### *Preliminary report of solar flare effects (s.f.e)*

Effects confirmed by ionospheric or solar observations are in italics.



1959 October 06d 14h 23m-14h 57m: Le Es  
Hu.—07d 14h 33m-14h 49m: Es (pt: MB).—  
10d 15h 04m-15h 08m: Vl Db Aq Md Fr MB  
(si: Es).—13d 09h 22m: IK Qu (s: Pi; bs: Bn).—  
14d 08h 57m: Ma Ks Mc? El (ssc: Cm; si: AA  
Hr; bs: Bn).—14d 12h 35m: Pa Va (si: Es Pi  
Hr; bps: Vl).—14d 19h 18m: Pi.—19d 11h 12m:  
Db Fr (si: Es).—20d 14h 40m: Hu.—28d 03h  
16m: Mb? Ka? Ky?

1959 November 09d 09h 33m: Ks (ssc: Cm).—  
14d 10h 30m: Mc? El.—16d 08h 09m: Qu Bi El  
(si: Ks Mc; b: Bn).—19d 14h 30m: Wi (bp: Es  
Db).—28d 12h 18m-13h 30m: Le.—30d 02h  
48m: Mb Ka Ky Hn Gn Am (ssc: Gu PM Tn To).

1959 December 02d 12h 46m: Le Es Wi M  
Eb Ta Pa Lr El Tn Va (si: MB Mc Bi Hr;  
Bn).—03d 18h 37m: Hu.—05d 12h 11m-13h 00  
Le Es.—06d 08h 07m: Db (si: MB).—08d  
59m: Lr.—11d 09h 11m: IK.—11d 15h 10  
Hu.—12d 11h 33m: Ma? Db (si: IK Mc  
14d 15h 31m: Hu.—16d 19h 12m: Hu.—17d  
01m: Hu.—21d 16h 25m: Hu.—21d 19h 44  
Hu Tw?.—22d 11h 52m: Fr.

COMMITTEE ON RAPID VARIATIONS AND  
EARTH CURRENTS

A. ROMAÑA, *Chairman*, Observatorio del Ebro  
Tortosa, Spain

# SELECTED GEOMAGNETIC AND SOLAR DATA

*Kp, Ci, Cp, Ap, K<sub>Fr</sub>, Rz, and Selected Days*  
May 1960

| Day <sup>1</sup> | 3-hr Range Indices <i>Kp</i> <sup>2</sup> |    |    |    |    |    |    |    |     | Prel. <sup>3</sup><br><i>Ci</i> | <i>Cp</i> <sup>4</sup> | <i>Ap</i> <sup>5</sup> | 3-hr Range Indices <i>K<sub>Fr</sub></i> <sup>6</sup> |     | Prov. <sup>7</sup><br><i>Rz</i> |
|------------------|---|----|----|----|----|----|----|----|-----|---------------------------------|------------------------|------------------------|---|-----|---------------------------------|
|                  | 1   | 2  | 3  | 4  | 5  | 6  | 7  | 8  | Sum |                                 |                        |                        | Values  | Sum |                                 |
| 1 D              | 7-  | 6+ | 6- | 4+ | 4o | 3+ | 3o | 4o | 37+ | 1.4                             | 1.6                    | 49                     | 6554 3234   | 32  | 97                              |
| 2                | 3-  | 4+ | 3o | 3o | 3- | 3o | 3o | 2o | 24- | 0.7                             | 0.9                    | 15                     | 2423 3233   | 22  | 97                              |
| 3 q              | 2-  | 2o | 3- | 3o | 2+ | 2o | 2- | 1- | 16o | 0.3                             | 0.4                    | 8                      | 2233 2220   | 16  | 102                             |
| 4 Q              | 1o  | 0+ | 1- | 1+ | 1o | 1o | 2+ | 3- | 10+ | 0.2                             | 0.2                    | 5                      | 0001 1233   | 10  | 96                              |
| 5 q              | 2-  | 2o | 3- | 2o | 2- | 1+ | 3+ | 4- | 18+ | 0.5                             | 0.6                    | 10                     | 2132 1144   | 18  | 87                              |
| 6 D              | 3+  | 4- | 4- | 5o | 4o | 5+ | 7o | 7+ | 39+ | 1.6                             | 1.7                    | 60                     | 4444 3457   | 35  | 93                              |
| 7 D              | 6-  | 6o | 4o | 5o | 6o | 5o | 4+ | 5+ | 41+ | 1.5                             | 1.6                    | 55                     | 5644 4444   | 35  | 133                             |
| 8 D              | 3+  | 6+ | 7- | 7+ | 8o | 8+ | 6+ | 7- | 53o | 1.9                             | 1.9                    | 128                    | 4657 6666   | 46  | 143                             |
| 9                | 3o  | 4- | 3o | 4o | 3+ | 3o | 2o | 2- | 24- | 1.0                             | 0.9                    | 16                     | 3434 3222   | 23  | 142                             |
| 0                | 2+  | 2o | 2+ | 2+ | 3- | 2+ | 3o | 4- | 21- | 0.7                             | 0.7                    | 12                     | 2233 3234   | 22  | 149                             |
| 1                | 4-  | 7- | 6+ | 4o | 3+ | 3o | 4- | 4o | 35- | 1.3                             | 1.5                    | 42                     | 4654 2245   | 32  | 147                             |
| 2                | 2o  | 4- | 4+ | 4o | 4+ | 4- | 2- | 2+ | 26o | 1.0                             | 1.0                    | 20                     | 2343 4323   | 24  | 127                             |
| 3 q              | 3-  | 2+ | 2- | 2o | 2+ | 3- | 3+ | 3+ | 20+ | 0.7                             | 0.7                    | 11                     | 2222 2243   | 19  | 135                             |
| 4                | 3o  | 3o | 4o | 3+ | 2o | 3- | 1+ | 2- | 21o | 0.7                             | 0.8                    | 13                     | 3455 2221   | 24  | 105                             |
| 5 q              | 0o  | 2+ | 3- | 4- | 3+ | 2o | 2- | 1- | 16+ | 0.6                             | 0.5                    | 10                     | 0232 2221   | 14  | 85                              |
| 6                | 1+  | 2- | 2o | 2o | 5+ | 6+ | 6+ | 6- | 31- | 1.3                             | 1.5                    | 42                     | 1322 5566   | 30  | 101                             |
| 7                | 3o  | 2o | 4- | 4- | 4- | 3- | 1+ | 1+ | 21+ | 0.9                             | 0.8                    | 14                     | 3244 4222   | 23  | 114                             |
| 8 Q              | 2-  | 2o | 2o | 1o | 1o | 2- | 1+ | 1+ | 12o | 0.2                             | 0.2                    | 6                      | 3321 1112   | 14  | 106                             |
| 9 Q              | 2o  | 2- | 2+ | 2- | 1- | 1- | 1o | 1+ | 11+ | 0.2                             | 0.2                    | 5                      | 2232 1011   | 12  | 108                             |
| 20 Q             | 1-  | 2+ | 0+ | 1- | 1- | 1o | 0+ | 1o | 7o  | 0.1                             | 0.1                    | 4                      | 1301 1111   | 09  | 115                             |
| 21 q             | 2-  | 1- | 1+ | 2- | 3- | 1+ | 1+ | 1- | 11+ | 0.2                             | 0.3                    | 6                      | 1122 3121   | 13  | 100                             |
| 22 Q             | 0+  | 1o | 0+ | 1- | 3- | 2- | 1+ | 2- | 10- | 0.3                             | 0.2                    | 5                      | 0001 3322   | 11  | 112                             |
| 23               | 1o  | 1o | 1+ | 2o | 6- | 5- | 5+ | 4o | 25o | 1.1                             | 1.2                    | 26                     | 1121 5464   | 24  | 125                             |
| 24               | 4-  | 5- | 6- | 4- | 4- | 3+ | 4+ | 4o | 33o | 1.2                             | 1.3                    | 31                     | 3453 3334   | 28  | 147                             |
| 25               | 4o  | 3+ | 3- | 3- | 3+ | 4- | 4- | 3+ | 27- | 1.1                             | 1.0                    | 19                     | 4423 3333   | 25  | 148                             |
| 26               | 4+  | 4o | 3+ | 3- | 2+ | 3o | 3o | 4o | 27- | 1.1                             | 1.0                    | 19                     | 4442 2234   | 25  | 130                             |
| 27               | 5o  | 4+ | 2o | 1+ | 2o | 2- | 3- | 3- | 22- | 0.9                             | 0.9                    | 16                     | 4311 2234   | 20  | 148                             |
| 28               | 2+  | 2- | 2o | 1+ | 1o | 1+ | 6o | 4+ | 20o | 0.9                             | 1.0                    | 18                     | 4222 1164   | 22  | 142                             |
| 29 D             | 8-  | 4+ | 4o | 3+ | 5o | 5- | 5+ | 4+ | 39- | 1.6                             | 1.6                    | 54                     | 7443 4455   | 36  | 138                             |
| 30               | 3-  | 2+ | 4- | 4o | 4+ | 4o | 2- | 2o | 25- | 1.1                             | 1.0                    | 18                     | 3344 4313   | 25  | 121                             |
| 31               | 2o  | 2+ | 2+ | 2+ | 3o | 3o | 3o | 4- | 22- | 0.9                             | 0.7                    | 13                     | 2333 3244   | 24  | 111                             |
| Means:           |   |    |    |    |    |    |    |    |     | 0.88                            | 0.90                   | 24                     |   |     | 119.5                           |
| No. of days:     |   |    |    |    |    |    |    |    |     | 31                              | 31                     | 31                     |   |     | 31                              |

- Five quiet days (Q), ten quiet days (Q or q), five disturbed days (D) selected by Committee on Characterization of Magnetic Disturbances, J. Veldkamp, Kon. Nederlandsch Meteorologisch Instituut, DeBilt, Holland.
- Geomagnetic planetary 3-hr-range indices *Kp* prepared by Committee on Characterization of Magnetic Disturbances, J. Bartels, Chairman, University, Göttingen, Germany.
- Preliminary magnetic character figures, *Ci*, prepared by J. Veldkamp.
- Magnetic character figures, *Cp*, prepared by J. Bartels.
- Average amplitudes *Ap* (unit 2γ), prepared by J. Bartels.
- Fredericksburg 3-hr-range indices *K* (*K9* = 500γ); scale values of variometers in γ/mm: *D* = 2.7, *Y* = 2.5, *Z* = 3.0; prepared by Robert E. Gebhardt, Observer-in-Charge, Fredericksburg Magnetic Observatory, Corbin, Virginia.
- Provisional sunspot numbers (dependent on observations at Zurich Observatory and its stations at Locarno and Arosa) prepared by M. Waldmeier, Swiss Federal Observatory, Zurich, Switzerland.

## Letters to the Editor

### Angular Motion of the Spin Axis of the Tiros I Meteorological Satellite Due to Magnetic and Gravitational Torques

WILLIAM R. BANDEEN

*Goddard Space Flight Center  
National Aeronautics and Space Administration  
Washington, D. C.*

AND

WARREN P. MANGER

*Astro-Electronic Products Division, Radio Corporation of America  
Princeton, New Jersey*

The Tiros I meteorological satellite was injected into orbit about the earth on April 1, 1960, at 11h 52m GMT. The purpose of the satellite was to photograph cloud patterns and distribution over the earth [Stroud, 1960].

In order to properly command the satellite from the ground to take direct or remote pictures, and in order to analyze these pictures when they are telemetered to the ground, the time, spatial coordinates, and spin-axis attitude of the satellite must be known. Ephemerides based on tracking data from the worldwide Minitrack Station Network and published by the NASA Space Computing Center determine accurately the spatial coordinates as a function of time. It was thought that, except for small perturbations, the spin-axis attitude would be space-stabilized for the nominal 3-month life of the satellite.

The spin rate of Tiros I at injection was 10.0 rpm. The axes of the two television cameras were parallel to the spin axis but pointing in the opposite direction (Fig. 1). A tuned energy-absorption mass device was installed in the satellite to damp out the initial free precession or nutation after separation from the third-stage rocket. It apparently has worked very well. The time of injection and the trajectory of the launch vehicle determined that the coordinates of the satellite spin vector on the celestial sphere should have been the following: declination,  $+19.8^\circ$ ; right ascension (RA),  $+58.6^\circ$  (Fig. 1).

After several days of picture taking, however,

it became apparent from analyses of photographs showing the horizon or identifiable landmarks that the direction of the spin vector (and hence, the camera axes) was not fixed but moving southward by as much at  $3^\circ$  to  $5^\circ$  per day (Fig. 2). On April 23, an analysis of photographs indicated that the spin axis had reached its southernmost declination, namely,  $-30^\circ$ . At this time its RA had increased to  $+69.0^\circ$ . Several days later it was determined from photographs that the spin vector was moving northward again and its RA increasing (eastward) at a greater rate than previously.

The Tiros I orbit is inclined to the equator at  $48.4^\circ$  and is nearly circular, having an apogee of 466 and a perigee of 431 statute miles. Its period is 99.24 minutes. Owing to a torque exerted on the orbit by the earth's bulge, the orbit regresses (westward) around the equator  $4.547^\circ/\text{day}$  (Fig. 1). It is apparent from Figure 1 that if the spin vector were to remain fixed in space, the westward motion of the orbit would cause an angle to develop between the orbital plane and the spin axis. Since Tiros I has the shape of a short cylinder (19 inches high and 42 inches in diameter), the moment of inertia about its spin axis is larger than the other principal moments of inertia. Hence, when the angle between the orbital plane and the spin axis is greater than zero and less than  $90^\circ$ , a torque due to differential gravity (similar to that exerted by the sun and moon on the earth, causing its precession) will be exerted on the satellite. On investigation

is torque was rejected as the primary torque using the observed forced precession because it could have caused the spin vector to move northward instead of in the observed southerly direction.

The minimum spin vector declination of  $-30^\circ$  was reached about April 23. At this time the orbit had regressed  $100^\circ$  and the unit vector  $\mathbf{n}$ , normal to the orbital plane, lay in the same meridional plane as the unit vector  $\mathbf{s}$  along the satellite spin axis. Shortly thereafter when  $\mathbf{n}$  developed a westerly component with respect to  $\mathbf{s}$  and the satellite spin-axis declination was observed moving northward, the possibility that magnetic couple was the primary force became apparent, and it became the focal point of investigation.

The outcome of this investigation has indicated that the angular motion of the spin axis of Tiros can be explained quite well by considering two torques: a primary torque caused by the interaction of a magnetic dipole along the satel-

lite spin axis (caused by residual permanent magnetism of ferromagnetic materials or by closed current loops in the instrumentation or by a combination of the two) with the earth's magnetic field, and a secondary torque caused by differential gravity in the earth's gravitational field.

The equation of motion expressing the action of these torques is

$$\frac{d\mathbf{s}}{dt} = \frac{1}{2}\epsilon(\mathbf{s} \cdot \mathbf{n})(\mathbf{s} \times \mathbf{n}) + \mu(\mathbf{b} \times \mathbf{s}) \quad (1)$$

where

$$\epsilon = \frac{3\omega_0^2}{\omega_s} \left( \frac{I - J}{I} \right) \quad \mu = \frac{M}{I\omega_s} \left( \frac{V_0}{R^3} \right)$$

and where

$\mathbf{s}$  = unit vector along the spin axis.

$\omega_0$  = angular velocity of the orbital radius vector.

$\omega_s$  = angular velocity of the satellite around its spin axis.

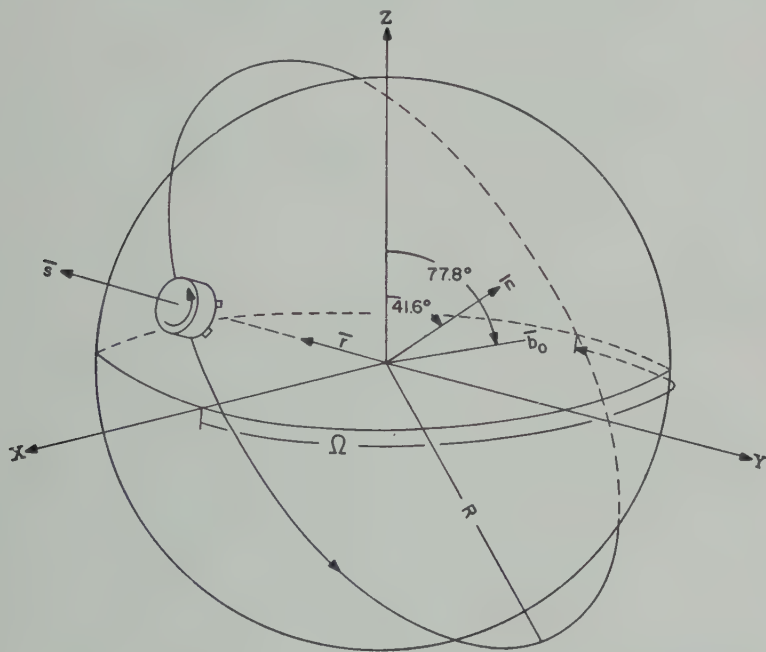


Fig. 1. Tiros I at injection, 1152 GMT, April 1, 1960.  $X$ ,  $Y$ , and  $Z$  are space coordinates with origin at the center of the earth and with the  $X$ - $Y$  plane in the equatorial plane of the earth.  $\mathbf{s}$  is the unit spin vector of the satellite. The astronomical coordinates of  $\mathbf{s}$  are declination,  $+19.8^\circ$ ; right ascension,  $+58.6^\circ$ .  $\Omega$  is the angle from the  $X$  axis to the orbital ascending node. The orbital nodes regress at the rate of  $-4.547^\circ \text{ day}^{-1}$  (westward).  $\mathbf{n}$  is the unit vector normal to the orbit.  $\mathbf{b}_0$  is the mean normalized magnetic dipole field vector appearing to the satellite in one orbit.  $\mathbf{r}$  is the unit vector from the center of the earth toward the satellite.  $R$  is the orbital radius.



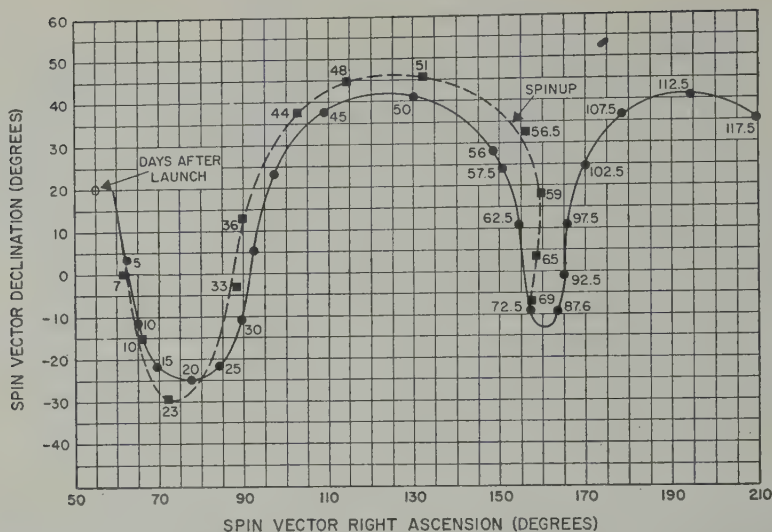


Fig. 2. Observed motion of the Tiros I spin vector based on an analysis of photographs (dashed line) compared with the theoretical motion based on the effects of a magnetic dipole moment along the spin axis and differential gravity (solid line). Declination is +north and -south of the celestial equator. Right ascension is +east of the vernal equinox along the celestial equator. Increased stability of both the theoretical and observed motion is seen after spinup on May 27 (day 56 after launch). The last picture with clearly identifiable landmarks was received on June 9 (day 69 after launch).

$I$  = moment of inertia around the spin axis.

$J$  = transverse moments of inertia.

$\mathbf{n}$  = unit vector normal to the orbit.

$M$  = magnetic dipole moment along the spin axis.

$V_0$  = magnetic constant for a dipole at the center of the earth.

$R$  = orbital radius.

$\mathbf{b}$  = normalized magnetic dipole field vector for the earth's field

$$\mathbf{b} = \mathbf{i}(-3r_x r_z) + \mathbf{j}(-3r_y r_z) + \mathbf{k}(1 - 3r_z^2)$$

where  $r_x$ ,  $r_y$ , and  $r_z$  are components along the space  $x$ ,  $y$  and  $z$  axes, respectively, of the unit vector  $\mathbf{r}$  from the center of the earth to the satellite.

Representative values of these parameters for Tiros I are the following:

$$\omega_0 = 0.001055 \text{ rad sec}^{-1}.$$

$$\omega_s = 1.047 \text{ rad sec}^{-1}.$$

$$I = 170,000 \text{ kg cm}^2.$$

$$J = 124,100 \text{ kg cm}^2.$$

$$M = 896 \text{ dyne cm gauss}^{-1}$$

(equivalent of 1 ampere flowing in 1 turn of wire around the satellite base plate).

$$\frac{V_0}{R^3} = 0.19 \text{ gauss}$$

If the above values are substituted in equation 1 we have:

$$\begin{aligned} d\mathbf{s}/dt = & \frac{1}{2}(8.6 \times 10^{-7})(\mathbf{s} \cdot \mathbf{n})(\mathbf{s} \times \mathbf{n}) \\ & + (9.6 \times 10^{-7})(\mathbf{b} \times \mathbf{s}) \end{aligned}$$

where  $d\mathbf{s}/dt$  is in radians per second.

Equation 2 was programmed on an electronic computer using the initial injection conditions of Tiros I on April 1, 1152 GMT. The results are plotted together with the observed motion of the spin axis in Figure 2, showing excellent agreement. In this computer run a time function for  $\omega_s$  was introduced reflecting its observed decrement from 10.0 rpm on April 1 to 9.4 rpm on May 27. (The decay in spin is due to a very small torque caused by eddy currents generated in the spinning satellite by the earth's magnetic field.) On May 27 a pair of spinup rockets was fired at 2133 GMT, increasing  $\omega_s$  to 12.0 rpm. As is seen from equation 1, greater stability is effected at higher values of  $\omega_s$ , and this increased stability can be seen in the plot of the spinup (Fig. 2).

In this work to date, a theoretical dipole field has been assumed for the earth's magnetic field. For this field the mean magnetic dipole field vector  $\mathbf{b}_0$  appearing to a satellite in a Tiros I type of orbit lies in the  $\mathbf{n}$ -earth axis plane at a latitude of  $77.8^\circ$  (Fig. 1). Further investigations of the theoretical aspects of this problem are now under way, including the use of a more realistic magnetic field. A detailed paper will be published in the near future.

The Tiros II meteorological satellite, soon to be launched, will carry a radiation experiment as well as two television cameras similar to those in Tiros I. It is especially important for the radiation experiment that the angle between the sun and the satellite spin axis not exceed  $35^\circ$ .

It is planned, therefore, to include a variable closed current loop in Tiros II to be commanded from the ground for the purpose of controlling the spin-axis attitude from the ground according to the requirements of the experiment.

*Acknowledgment.* We should like to express our thanks to I. Ruff and S. Twomey of the U. S. Weather Bureau, and to R. Goerss of RCA, all of whom made contributions to the investigation.

#### REFERENCES

- Stroud, W. G., Initial results of the Tiros I meteorological satellite, *J. Geophys. Research*, **65**, 1643-1644, 1960.

(Received July 13, 1960.)

# A Method for Measuring Temperature Directly in the Upper Atmosphere with a Rocket-Borne Magnetic Mass Spectrometer

C. Y. JOHNSON, J. H. HOFFMAN, J. M. YOUNG, AND J. C. HOLMES

*U. S. Naval Research Laboratory  
Washington 25, D. C.*

This note describes briefly a method for measuring the temperature of the atmosphere above 100 km directly by means of a rocket-borne magnetic mass spectrometer. Previously, above this altitude, temperature has been a derived parameter obtained from pressure or density data.

When a vehicle passes through a gas, the number of molecules passing through a collimating entrance system mounted on the vehicle is a function of the velocity of the vehicle relative to the gas, the attitude of the entrance system with respect to the vehicle trajectory, and the temperature of the gas as derived from an assumed velocity distribution. If the velocity and attitude of the vehicle are known, and the relative numbers of molecules of a specific atmospheric constituent passing through the entrance system can be measured, the temperature of that constituent can be determined.

On June 27, 1960, an Aerobee-Hi rocket, NRL-58, was launched at the White Sands Missile Range, New Mexico. The rocket reached a peak altitude of 220 km with a 210-pound payload, which included a single focusing Nier type 60° magnetic mass spectrometer having a 1½-inch-radius magnetic analyzer and a special entrance system mounted on the side of the rocket. The entrance system as shown in Figure 1 con-

sisted of two 8° half-angle cones separated by a plate containing a hole concentric with the axes. The first cone opened toward the atmosphere; the second, into the mass spectrometer ionization chamber. The region between the cone and plate was 'purged' by a 1 liter/sec titanium getter ion pump (Vacion); the spectrometer tube and second cone were purged by a 5 liter/sec Vacion pump. Gas molecules entering from the angle subtended by the walls passed into the mass spectrometer unperturbed, but the great majority of those without this 'acceptance' cone were either rejected by the outer cone or removed from the stream by the pumps. Before the flight, the outer cone was covered with a spring-loaded cap sealed to the system by a glass ring. This ring was broken by two squib-activated hammers at 150 km altitude on the ascent of the rocket. Signals were then recorded from 150 km to peak altitude of 220 km and down to 80 km on the descent, at which altitude the instrument was severed from the rocket.

The mass spectrometer ion source was of the conventional electron bombardment type. The collector assembly consisted of two slits, placed at the normal focal point of the magnetic analyzer, the other in a position to collect mass 40 ions when mass 28 ions are focused at the first slit. Thus both N<sub>2</sub> and A ions were collected simultaneously. Each collector slit was followed by a ten-stage BeCu electron multiplier having a gain of about 10<sup>4</sup> at 2500 v. Each multiplier output was suitably amplified and telemetered. The instrument was scanned, but remained focused on masses 28 and 40 throughout flight.

In order to minimize background ion current from residual gas and to minimize back diffusion of gas through the entrance system, both the mass spectrometer tube and the entrance

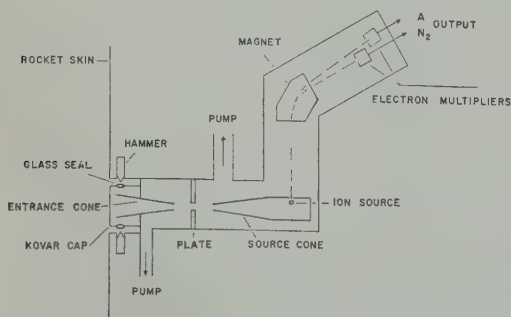


Fig. 1.

n were pumped continuously before and during flight by the two Vac-ion pumps. For several weeks before the flight the pressure in the system as determined by the Vac-ion pump current was less than  $10^{-8}$  mm Hg. (It read zero on the Vac-ion indicator meter.)

The purpose of this experiment, as previously stated, was to determine the temperature of the atmosphere by measuring the flow through an entrance system of a specific atmospheric constituent. The mass spectrometer then served as a detector which determined the number of molecules passing through the entrance system. Furthermore, it sorted molecules according to mass, thus enabling a temperature determination to be made from an assumed velocity distribution. Two gases, molecular nitrogen and argon, were chosen so that two independent temperature measurements could be made simultaneously. Because of the roll of the rocket about its axis, and the fact that the rocket was traveling at an appreciable fraction of the mean molecular velocities of the atmospheric gas particles, the spectrometer ion beams were modulated. When the entrance system pointed forward along the rocket trajectory, an increase in spectrometer ion beam intensity was observed due to 'ram.' As the rocket rolled and the entrance system pointed backward, the ion beam intensities were decreased by 'suction.'

Thus the ion beam intensity modulation was a function of the attitude and velocity of the entrance system and the velocity distribution of the atmospheric molecules, or their temperature. Rocket attitude was known from the optical and magnetic aspect devices carried in the rocket. Radar tracking of the rocket provided the required velocity data. Thus, in principle, from the ratio of the ion beam amplitudes at various rocket attitudes the temperature of the gas through which the rocket is passing can be determined. In addition to the temperature measurement, the ratio of the molecular nitrogen to argon ion beam amplitudes versus altitude can yield a diffusive separation measurement of these two gases.

In the experiment of June 27, 1960, after the spectrometer entrance system was opened to the upper atmosphere the Vac-ion pump monitor indicated a pressure of approximately  $10^{-4}$  mm Hg at the mouth of the entrance cone. Because of this constant high-pressure background, assumed to be from either a gas leak in the nose cone skin or surface outgassing, the percentage modulation recorded was reduced and seldom exceeded 11 per cent. Therefore, from this first attempt an accurate temperature profile will not be obtained. Additional flights are planned.

(Received July 14, 1960.)



Outer Radiation Belt and Solar Proton  
Observations with Explorer VII  
during March–April 1960<sup>1</sup>

JAMES A. VAN ALLEN AND WEI CHING LIN

Department of Physics and Astronomy, State University of Iowa  
Iowa City, Iowa

*Introduction.* The IGY composite satellite Explorer VII (1959*t*), launched on October 13, 1959, includes an instrument prepared by this laboratory [Ludwig and Whelpley, 1960] for a 1-year comprehensive study of (a) the lower parts of the inner and outer radiation belts, (b) the primary cosmic-ray intensity near the earth, and (c) the arrival of solar protons. Table 1 summarizes the properties of the two Geiger tube detectors in this instrument.

As of April 1, 1960, the orbit of Explorer VII was characterized by the following: perigee altitude 565 km, apogee altitude 1096 km, inclination 50.5°, anomalistic period 101.3 minutes. Typically, about five passes per day, each giving some 5 to 20 minutes of substantially continuous data, are received at the Iowa City station. The track of the satellite during the most favorable one of each day's passes is from

northern California, across northwestern United States, southern Canada (nearly to the southern tip of Hudson Bay), and Newfoundland, terminating in mid-Atlantic. In the study of the effects at high magnetic latitude, observations along such portions of the orbit are the most valuable of any obtained with the present satellite. The next most advantageous portions are to the south of Australia. All the data of the present report were derived from receptions at Iowa City and at the NASA stations at Bloss Point, Maryland, and Woomera, Australia. The positional coordinates of the satellite are determined by the NASA tracking chain and the Space Computing Center of the Goddard Space Flight Laboratory. We are indebted to the agency for the prompt receipt of the tabular ephemeris at 1-minute intervals of time. The present note is concerned with the period of the latter part of March to mid-April 1960, being published in advance of a more complete

<sup>1</sup> Preliminary report given at American Geophysical Union Symposium, April 29, 1960.

TABLE 1. Properties of SUI Detectors in Explorer VII

| Detector                      | $\epsilon G_0^*$    | Absorbers over 70%<br>of Solid Angle    | Approximate Detection Threshold |                                    |                        |
|-------------------------------|---------------------|---|---------------------------------|------------------------------------|------------------------|
|                               |                     |   | Proton                          | Electron,† ex-<br>trapolated range | X Rays<br>transmission |
| Anton type 112 Geiger<br>Tube | 7.2 cm <sup>2</sup> | Stainless steel 0.040 g/cm <sup>2</sup> | 30 Mev                          | 2.5 Mev                            | 80 kev                 |
| Scaling Factor: 128           |                     | Al 0.26                                 |                                 |                                    |                        |
|                               |                     | Pb 1.15                                 |                                 |                                    |                        |
|                               |                     | Mg 0.14                                 |                                 |                                    |                        |
| Anton type 302 Geiger<br>Tube | 0.54                | Stainless steel 0.40                    | 18                              | 1.1                                | 30                     |
| Scaling Factor: 2048          |                     | Mg 0.14                                 |                                 |                                    |                        |

\* Counting rate of tube for penetrating particles is equal to  $\epsilon G_0$  times omnidirectional intensity  $J_0$ .  
† For nonpenetrating electrons, sample experimental values of ratio of omnidirectional intensity to counting rate of 302 tube are:  $10^{13}$  at 14 kev;  $10^{12}$  at 17 kev;  $10^{11}$  at 20 kev;  $10^{10}$  at 26 kev;  $10^9$  at 34 kev; at 45 kev;  $10^7$  at 70 kev;  $2 \times 10^6$  at 105 kev. (L. A. Frank, private communication, April 1960. Experimental values obtained by electron bombardment of an arrangement similar to SUI package in Explorer VII)

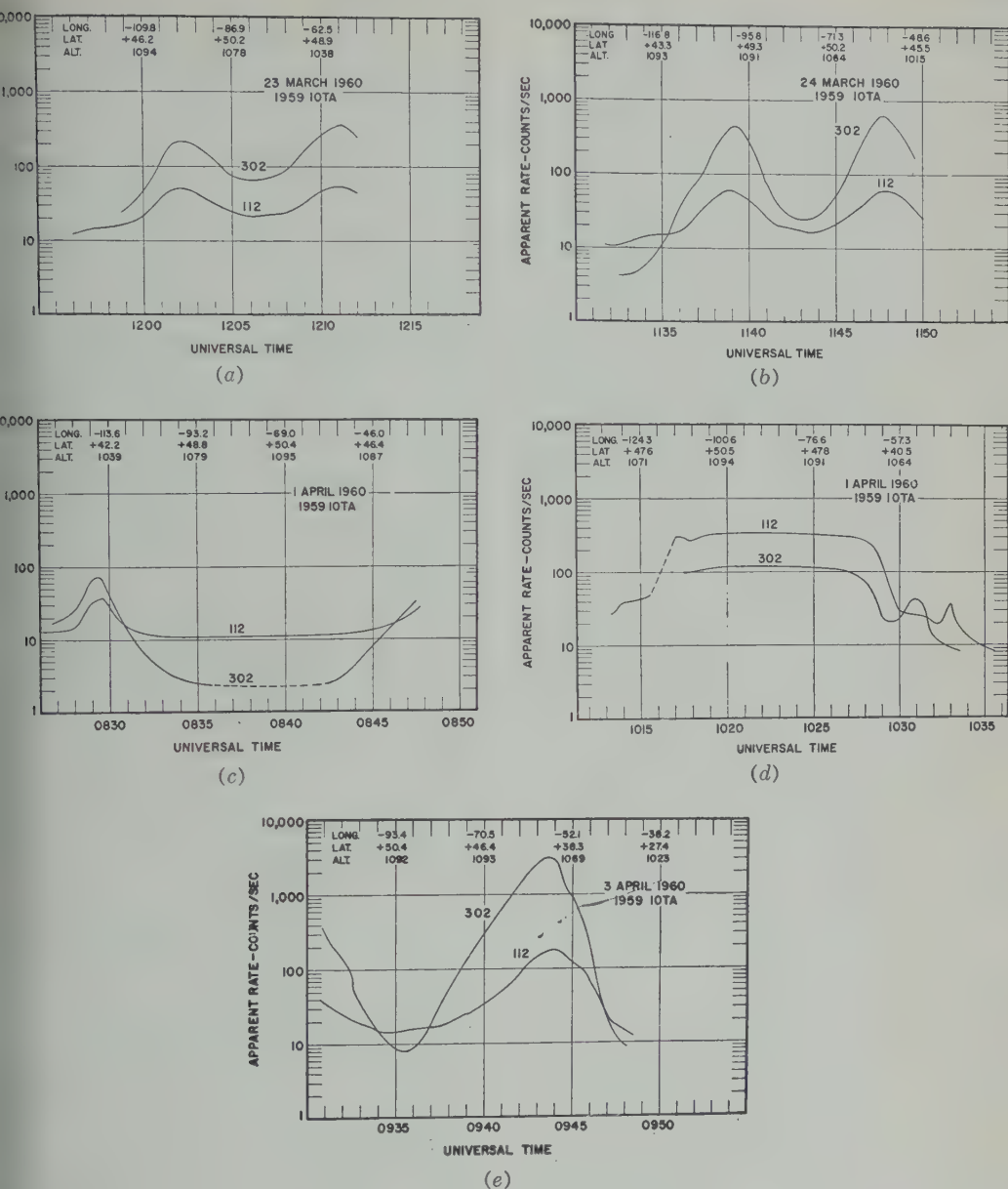


Fig. 1. A sequence of representative Explorer VII (1959) observations from the latter part of March to early April 1960, from the Iowa City receiving station.

ensive survey because of the geophysical events of special interest during the period mentioned and because of the valuable support of the related and simultaneous observations with Pioneer V, which was in the near astronomical vicinity of the earth but well beyond its magnetic influence during this period (radial dis-

tance of 0.036 astronomical unit, approximately toward the sun, on April 1).

*Observational data and discussion.* Figures 1a, b, c, d and e exhibit the counting-rate data from a series of representative passes. The plots for March 23 and 24 typify quiescent outer-zone conditions during the several weeks preceding

March 31. The one for 0826-0847 of April 1 shows a great depletion of the outer zone, its movement to much lower latitudes, and the nearly complete absence of soft radiation at high magnetic latitudes. The plot for 1013-1035 of April 1 shows the presence of an abnormal intensity of relatively penetrating radiation which is confined to magnetic latitudes exceeding  $57^\circ$ . This radiation is not conclusively identified by the present experiment. But when our data are taken in conjunction with the large body of knowledge about such events which has been developed during the past two years we feel justified in regarding the anomalous radiation as consisting of solar-emitted protons. Figure 1e exhibits the rapid recovery of the outer zone, its return to high latitudes, and the great enhancement of intensity that followed the events of March 31-April 1.

By plotting the rate of the 112 (the larger, more heavily shielded tube) against that of the 302 as the northern extent of the outer zone varies, we find it possible to make a reliable subtraction of the contribution of trapped (or other

soft) radiation to the counting rate of the 112 at high magnetic latitudes (over southern Canada). The temporal dependence of the rate due to the penetrating component, also corrected for artificial dead time of the circuit [Ludwig & Whelpley, 1960], is plotted in the upper part of Figure 2. The normal cosmic-ray rate during the period March 23-30 was  $14.7 \pm 0.2 \text{ sec}^{-1}$ , corresponding to  $J_0 = 2.0 \pm 0.1 (\text{cm}^2 \text{ sec})^{-1}$ . During the early morning of April 1 the rate was found to have declined to  $11.2 \pm 0.3$  (a decrease of 24 per cent). The onset of this Forbush decrease was missed in our data but was observed at about 1200 on March 31 by Carmichael & Steljes (data from neutron intensity monitor at Deep River, Canada,  $46^\circ 06' \text{N}$ ,  $77^\circ 30' \text{W}$ , mean pressure 1000 mb, private communication, April 1960); the decrease had reached its greatest value, 10 per cent, at the Deep River station about 0600 on April 1.

The following satellite information is available about the onset of the April 1 solar proton event: (1) As late as 0842 the corrected rate of the 112 did not exceed  $11.0 \text{ sec}^{-1}$  at high magnetic latitude (Fig. 1c, Iowa City station).

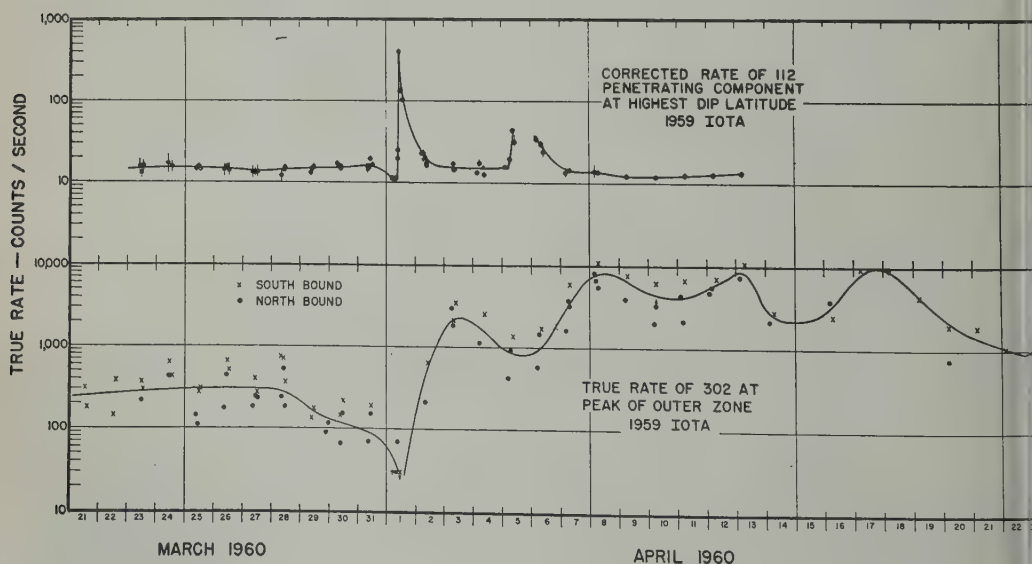


Fig. 2. Upper part shows the temporal dependence of the rate of the 112 Geiger tube after subtraction of soft radiation. The data are from the time periods of the high-latitude 'valley' of such curves as shown in Figures 1a-e (Iowa City and Woomera stations). Lower part shows the temporal dependence of the true rate of the 302 Geiger tube (lightly shielded) at the peaks of the outer zone observed at about 1000-km altitude over North America (Iowa City station). Scatter of points is presumably due to combination of short-term temporal variations, differences in geomagnetic altitude and longitude, and changing relative satellite orientation.



the next pertinent pass was observed from the Voomera station. The corrected rate rose from 3 to 35 in the interval 0933–0936, then varied regularly between 22 and 25, diminished rapidly and smoothly to 10 in the interval 0938–0940 as the satellite passed to lower latitudes, and finally continued the usual gradual cosmic-ray decline as it moved toward the equator. (3) On May 1019 (Fig. 1d) the rate of the 112 had reached the highest observed value during the event.

From these data we conclude that a substantial increase of the intensity of protons of  $E > 30$  Mev. had not occurred before 0933, but did occur during the period 0933–1019. H. Leinach (polar-cap absorption events of March 31 through May 13, 1960, Geophysical Institute, University of Alaska, private communication, June 21, 1960) reported the onset of polar-cap absorption at 0945 at Thule. The optical flare which is regarded most plausibly as being the source of the April 1 solar proton event started at 0845 and ended at 1222 (Stockholm). Thus, assuming that the emission of energetic particles by the flare began immediately, the differential time lag between the arrival at the earth of light and of particles was between 48 and 60 minutes. The total flight time of particles was therefore 56 to 68 minutes. For rectilinear flight from the sun to the earth (center to center), protons of the following kinetic energies have the following flight times: 10 Mev, 58 minutes; 20 Mev, 41 minutes; 30 Mev, 34 minutes; 40 Mev, 29 minutes. Hence the observed delay of about 1 hour suggests one of four conclusions: (1) the flight path was about twice the rectilinear one (owing to magnetic fields near the solar source and in interplanetary space); (2) there was a time lag of the order of 0.5 hour between the optical outburst and the emission of a substantial intensity of protons of  $E > 30$  Mev; (3) a distortion of the geomagnetic field diverted the particles from coming close to the earth (even at Thule) before about 0945, at which time a marked change in the field occurred (*Chinburg* [1960] does indeed report a marked reduction of  $F$  beginning at 1045); or (4) there was a composite of effects 1, 2, and 3.

The maximum observed value of the corrected counting rate ( $402 \text{ sec}^{-1}$ ) of the 112 occurred in

the period 1019–1028 (cf. Fig. 1d). The corresponding maximum value of the absolute omnidirectional intensity of solar protons was  $54 (\text{cm}^2 \text{ sec})^{-1}$  from the 112 data; the simultaneous value from the 302 data was  $210 (\text{cm}^2 \text{ sec})^{-1}$ . These two data indicate that the integral number-energy spectrum was rising between 30 and 18 Mev approximately as  $E^{-2.6}$ . The latitude dependence of intensity (1027–1030 of Fig. 1d) has been studied in detail and has been found quite inconsistent with a proton spectrum of this form. It is concluded tentatively that the 'normal' geomagnetic cutoff energies were greatly reduced by the geomagnetic perturbations of this period. According to R. J. Naumann (Army Ballistic Missile Agency, private communication, May 1960), the forward-projected axis of Explorer VII has pierced the celestial sphere at a substantially fixed point for nearly all its lifetime, namely at right ascension  $(305 \pm 5)^\circ$ , declination  $+(27 \pm 5)^\circ$ . Thus, our solar proton data over northern United States and southern Canada for the April 1 event were obtained while the zenith angle of the payload axis did not exceed  $50^\circ$  and was usually considerably less. Since the SUI instrument package was located on the extreme forward end of the (axially spinning) satellite, the thinly shielded portion of the solid angle of the Geiger tubes was more or less upward. During the south-of-Australia pass which yielded the two data points on the early part of the sharply rising side of the event, the zenith angle was in the range  $130^\circ$  to  $160^\circ$ .

The time width of the April 1 event at half-intensity was about 2 hours, and the decay of intensity was after the fashion of  $(\Delta t)^{-2.6}$  up to an elapsed time  $\Delta t$  as great as 70 hours.

The solar proton event of April 1 was the most intense of the some ten events observed by Explorer VII since October 13, 1959. In comparison with events of the IGY period, however, this event would be classed as one of intermediate intensity, being similar to the ones of August 16–17, 23–24, and 26–27, 1958, which were observed with similar equipment in Explorer IV [*Rothwell and McIlwain*, 1959].

Another distinct solar proton event had its onset at about 0615 on April 5 and disappeared into the background 2 to 3 days later. Unfortunately, no suitable observations with Explorer VII are available for the period 1200 on April 5



to 0400 on April 6. Hence the structure of this event is not known from our work.

The cosmic-ray intensity during the period April 9-13 continued to be depressed, though it was gradually recovering toward its normal value.

The detailed history of the outer radiation belt during late March and early April contains fluctuations of large magnitude. The simplest aspect of these fluctuations is shown in the lower part of Figure 2, where we have plotted the temporal dependence of the counting rate of the 302 (corrected for dead time) at the peaks of the outer zone as it is traversed at an altitude of about 1000 km (cf. Fig. 1*a-e*). The most noteworthy features of the temporal dependence are the almost complete disappearance of the outer zone (as seen at low altitudes) on April 1, its replenishment over the following 2 days to an intensity 10 times the quiescent value during the latter part of March, and its further replenishment during the following week to an intensity 30 times its quiescent value or at least 300 times its minimum value on April 1. Moreover, some of the passes exhibited a complex spatial structure of the outer zone during the period of rapid growth.

The magnetic storm which began at 0815 UT on March 31 was among the largest on record [Chapman and Bartels, 1940]. The total excursion of scalar  $F$  during the 4-day period March 31-April 3 was recorded by the SUI proton precession magnetometer ( $41^{\circ} 40'N$ ,  $91^{\circ} 32'W$ ) as 1149 gammas. And the greatest deviation from the quiet mean value was +692 gammas at 2239 UT on March 31 [Chinburg, 1960]. The general correlation of the radiation observations with magnetic activity during March 31-April 3 is obvious. An attempt to discern detailed relationships has not been particularly fruitful. Low-latitude auroras over a wide range of longitudes occurred on the night of March 31-April 1. It is probable that the solar plasma cloud which was responsible for the magnetic storm and the precipitation of charged particles into the atmosphere was observed in interplanetary space by Winckler, Arnoldy, and Hoffman [1960] with an Anton 302 Geiger tube and an ionization chamber in Pioneer V.

An important aspect of these latter observations is that at no time during the period March

31-April 6 did the net counting rate due to solar radiation as measured by the 302 in Pioneer exceed 3 counts/sec. Inasmuch as the shielding of this tube was similar to the nearly identical one in Explorer VII, it is seen that the outer belt of trapped particles developed over a 2-day period to contain several thousand times the intensity of electrons of energy greater than, say, 30 keV that was present in the original solar plasma cloud before its arrival at the earth. This fact provides strong evidence for the 'local' acceleration of trapped electrons within the geomagnetic field with a characteristic time of the order of 1 day. Our low-altitude measurements with Explorer VII do not, of course, permit this conclusion to be altogether rigorous since it is conceivable that the build-up of the intensity in the outer zone as seen at low altitudes might have been the result of the downward perturbation of the mirror points of particles in the outer zone which were previously unobservable at an altitude of about 1000 km. In a similar event of lesser magnitude on August 16-18, 1959, however, Arnoldy, Hoffman, and Winckler [1960] found with Explorer VI that the time history of the entire outer zone was similar to that observed by us at low altitude on the March-April 1960 occasion and similar to that found with Explorer IV during August-September 1958 by Rothwell and McIlwain [1960].

The passage of magnetized solar plasma through interplanetary space was observed during March 31-April 1 with a one-component magnetometer in Pioneer V by Coleman, Sone, Judge, and Smith [1960]. Fan, Meyer, and Simonson [1960], with a system of counters on Pioneer V, observed a sequence of events during March-April similar to those described herein. The Forbush decrease on March 31-April 1 as reported by the latter authors was some 28 per cent on the triple coincidence telescope and some 15 per cent on the single, central counter of the array. These results may be compared to our value of 24 per cent at high magnetic latitude, an altitude of about 1000 km and to Winckler's similar value. Thus it appears that (at least on this occasion) the full amplitude of an interplanetary Forbush decrease is observed at high latitudes near the earth. Viewed in another way it appears that the Forbush decrease on the

occasion had its full development at a distance of 0.036 astronomical unit from the earth and was neither augmented nor diminished in the near vicinity of the earth.

*Acknowledgments.* The research for this paper was assisted by the National Aeronautics and Space Administration and by the joint program of the Office of Naval Research and the Atomic Energy Commission.

## REFERENCES

- Arnoldy, R. L., R. A. Hoffman, and J. R. Winckler, Observations of the Van Allen radiation regions during August and September 1959, part 1, *J. Geophys. Research*, *65*, 1361-1376, 1960.
- Chapman, S., and J. Bartels, *Geomagnetism*, vol. I, pp. 328-335, Clarendon Press, Oxford, 1940.
- Chinburg, D. L., Great magnetic storm of March 31 to April 3, 1960, *J. Geophys. Research*, *65*, 2206-2208, 1960.
- Coleman, P. J., Jr., C. P. Sonett, D. L. Judge, and E. J. Smith, Some preliminary results of the Pioneer V magnetometer experiment, *J. Geophys. Research*, *65*, 1856-1857, 1960.
- Fan, C. Y., P. Meyer, and J. A. Simpson, Preliminary results from the space probe Pioneer V, *J. Geophys. Research*, *65*, 1862-1863, 1960.
- Ludwig, G. H., and W. A. Whelpley, Corpuscular radiation experiment of satellite 1959<sub>u</sub> (Explorer VII), *J. Geophys. Research*, *65*, 1119-1124, 1960.
- Rothwell, P., and C. McIlwain, Satellite observations of solar cosmic rays, *Nature*, *184*, 138-140, 1959.
- Rothwell, P., and C. E. McIlwain, Magnetic storms and the Van Allen radiation belts: Observations from satellite 1958<sub>e</sub> (Explorer IV), *J. Geophys. Research*, *65*, 799-806, 1960.
- Winckler, J. R., R. L. Arnoldy, and R. A. Hoffman, Solar cosmic rays and soft radiation observed at 5,000,000 kilometers from earth, *J. Geophys. Research*, *65*, 3004-3007, 1960.

(Received July 1, 1960.)

## Solar Cosmic Rays and Soft Radiation Observed at 5,000,000 Kilometers from Earth

R. L. ARNOLDY, R. A. HOFFMAN, AND J. R. WINCKLER

*School of Physics, University of Minnesota  
Minneapolis, Minnesota*

During the period March 27 to April 6, 1960, the integrating ionization chamber and Geiger counter in the space probe Pioneer V detected solar cosmic rays and some soft radiation effects associated with a high level of solar activity. The instruments are of the same type as used on Explorer VI earth satellite [Arnoldy, Hoffman, and Winckler, 1960]. At this time the space probe was about 5,000,000 km from the earth, approximately in the plane of the ecliptic, and located somewhat behind the sun-earth radius toward the sun. The solar activity was associated with McMath plage region 5615 [National Bureau of Standards, 1960], which crossed the solar meridian on April 1, 1960. The region was characterized by numerous flares of all sizes, large loops and surge prominences, and strong solar radio emission over a wide range of frequencies. On March 31 at 0800 UT, a severe geomagnetic storm began on the earth accompanied by major earth current disturbances, a complete blackout of the North Atlantic communications channel, and auroral displays. Also, at this time, a large Forbush decrease occurred in the galactic cosmic radiation. An intense series of balloon flights was made at Minneapolis during the period, and counting rate increases were observed at high altitude due to solar cosmic rays and auroral X rays. The earth satellite Explorer VII showed very substantial changes in the earth radiation belts and also detected the solar cosmic rays.

The principal results from Pioneer V may be summarized as follows:

1. On March 11, data were obtained as Pioneer V left the outer Van Allen radiation region. The values of ionization and count rates, and the extent of the outer zone, agree well with our previous measurements on Explorer VI during August 1959 [Arnoldy, Hoffman, and Winckler, 1960].

2. For the first 2 weeks the galactic rates in

space showed a slow increase, also seen on the sea-level neutron monitors and balloon flights. The ionization/count ratio during this time is 1.4 times minimum, and the omnidirectional flux measured by the Anton counter is  $4.5 \pm 0.1$  cm<sup>2</sup>·sec. The total ionization rate on the normalized IGY scale [Winckler, 1960] is 26 P/sec 10<sup>3</sup>, corresponding to 318 ion pairs/cm<sup>3</sup>·sec. This is 1.45 times the normalized balloon ionization rate at 7 g/cm<sup>2</sup> atmospheric depth over Minneapolis, measured simultaneously. Geiger counter statistics are  $\pm 3$  per cent, and ionization chamber statistics negligibly small.

3. Beginning on March 27, region 5615 became active. Numerous rate increases were observed in Pioneer V, which are numbered in Figure 1 and identified in Table 1. The background line for galactic cosmic rays in space (dotted lines, Fig. 1) is obtained by normalizing balloon and space-probe ion chambers in the period March 11 to 25 and after April 11. Most of the increases seem to be the low-energy solar cosmic rays of the type observed previously on many occasions [Anderson, Arnoldy, Hoffman, Peterson, and Winckler, 1959; Ney, Winckler, and Freier, 1959; Winckler and Bhavsar, 1960; Freier, Ney and Winckler 1959]. For the first time we have now observed these solar cosmic rays in space completely free of the environmental effects of the earth. The largest cosmic-ray bursts occurred on April 1, associated with a class 3B flare beginning at 0830 UT, and on April 4, associated with a major radio noise storm beginning approximately 1200 UT with no flare as yet reported.

4. A balloon flight at Minneapolis on April 1 showed the presence of the solar protons at energies up to at least 400 Mev. The flux is consistent with the space-probe results. A severe geomagnetic storm was present during the April 1 flare, and the altered geomagnetic cutoff energies are the likely cause of the presence of the

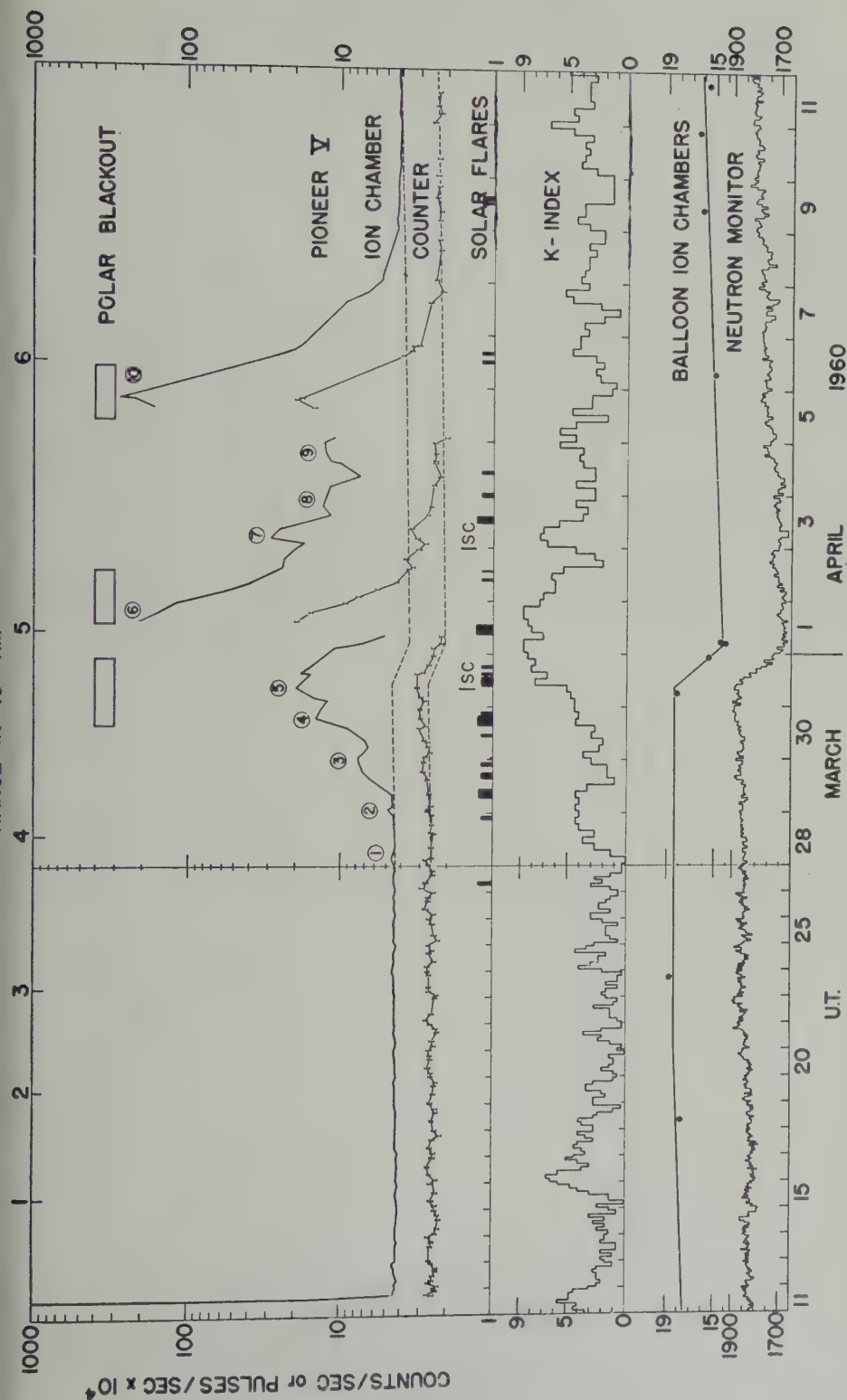


Fig. 1. Top to bottom: polar blackouts (courtesy H. Leinbach, College, Alaska); Pioneer V ion chamber Geiger counter, not normalized; solar flares, class 2 or greater; planetary  $k$  indices (J. Bartels, Göttingen); balloon total ionization at  $7 \text{ g/cm}^2$  atmospheric depth over Minneapolis on normalized scale adopted for IGY [Winckler, 1960]; sea-level neutron rate (courtesy H. Carmichael, Deep River, Ontario). Note expanded time scale after March 27. Ranges in kilometers from earth are approximate. Note: IGY standard ion-chamber rate = Pioneer V  $\times 60.7$ ; Anton counter omnidirectional area factor =  $0.50 \text{ cm}^2$ .



TABLE 1. Rate Increases Observed in Pioneer V

| Event No. | Ions/Count<br>× Minimum | Tentative<br>Identification | Excess Flux,<br>particles/cm <sup>2</sup> ·sec | Remarks  |
|-----------|-------------------------|-----------------------------|--|--|
| 1         | 1.5                     | Protons                     | 0.2  | Low accuracy   |
| 2         | 2.3                     | Protons                     | 0.2  | Low accuracy   |
| 3         | 9.5                     | Protons                     | 0.4  | Successive addition of protons                         |
| 4         | 17.7                    | Protons + X rays            | ...  | events and soft radiation                              |
| 5         | 19.5                    | Protons + X rays            | ...  | during magnetic storm                                  |
| 6         | 7.3                     | Protons                     | 35   | Source: flare of 0845 UT,<br>April 1                   |
| 7         | 14.6                    | Protons + X rays            | 2  | Total: see 7b below                                    |
| 8         | 13.3                    | Protons + X rays            | 1  |  |
| 9         | 20.6                    | X rays                      | ...  |  |
| 10        | 10.8                    | Protons                     | 32   | No reported flare<br>Noise storm ≈ 1200 UT,<br>April 5 |
| 7b        | 19.5                    | X rays                      | ...  | Excess over protons<br>Associated with S.C.            |

protons as far south, geomagnetically, as Minneapolis. Balloon observations failed to show any solar cosmic rays for the April 5 event.

5. From the balloon flight cosmic-ray telescope measurements for the April 1 event, a proton differential energy spectrum of the form  $E^{-n}$  is obtained, where  $3.5 < n < 4$ , for energies greater than 130 Mev (Thomas May, Andrew Masley, and JRW, University of Minnesota, private communication). These observations were made between 0950 and 1105 UT. For the two major cosmic-ray events, the observed ion/count ratio times minimum for the Pioneer V detectors is from 7 to 10, the April 1 observation being 5 hours after the commencement of the flare. Assuming that the balloon telescope spectrum remains constant in time, the power spectrum must be cut off at 30 Mev to obtain the above ratios. The peak measured dosage for the April 1 event was  $2.6 \times 10^{-2}$  r/hr. The decrease of the particles in space, timed from the associated event, follows the law  $I = I_0 T^{-1.9}$  with  $T$  measured in multiples of 1 hour. This decay is, within error, the same as for several known events including the great flare of February 23, 1956.

6. Polar blackouts were observed by Leinbach at College, Alaska, for three of the events as shown in Figure 1 (private communication).

7. Soft radiation, probably bremsstrahlung X rays similar to those seen in the outer radiation zone [Arnoldy, Hoffman, Winckler, 1960], or in auroras [Winckler, 1960], was apparently detected on March 31 (peak 5, Fig. 1) during the

intense magnetic storm and Forbush event, and again on April 3 (peak 7, Fig. 1). This radiation is identified by the high ion/count ratio, approaching 20 times minimum. Although such ratios may be obtained for solar protons with energies near the equipment cutoff (approximately 20 Mev), in the major proton events (peaks 6 and 10) ratios of about 10 were obtained. The sudden increase in ratio, during decay of a proton event, and closely associated with the sudden commencement (e.g., peak 5) tentatively identifies the increase as electron bremsstrahlung. The ion-chamber and Geiger counter readings on peak 5, assuming electrons correspond to a 50-kev electron flux of approximately  $10^7$ /cm<sup>2</sup>·sec as an upper limit striking the outer shell of Pioneer V. These electrons in free space must be the extreme energetic part of a solar plasma stream associated with the strong geomagnetic storm on March 31 and again with the weaker event on April 3.

It is of great importance to relate the above results to simultaneous measurements made with Explorer VII earth satellite and described in the accompanying letter by Van Allen and Lin. Two significant facts emerge from this comparison:

1. The solar proton fluxes measured by Antares counters in Pioneer V, and in Explorer VII during a high-latitude pass over southern Canada are in reasonable agreement during the April 1 cosmic-ray event.

2. Explorer VII observations showed a large initial decrease in the outer radiation zone

000-km altitude on March 31 followed by a very large increase during the following week.<sup>1</sup> The increase in electron flux may have reached  $10^{11}/\text{cm}^2 \cdot \text{sec}$  in the 50-kev range. At no time in space was a flux of similar energy electrons greater than  $10^7/\text{cm}^2 \cdot \text{sec}$  observed. If the energetic part of the outer radiation zone is caused by direct solar injection, a reasonable requirement is the equality of fluxes outside and inside the geomagnetic field. Since this is not observed, we conclude that the energetic part of the outer zone, which is the usual component measured so far in space experiments, arises from a local acceleration process within the confines of the geomagnetic field itself.

*Acknowledgments.* The space program was supported by National Aeronautics and Space Administration, contract NASw-56. Balloon flights were sponsored by the National Science Foundation.

<sup>1</sup> Similar behavior was observed in the outer zone at large distances by Explorer VI in August 1959 [Arnoldy, Hoffman, and Winckler, 1960].

## REFERENCES

- Anderson, K. A., R. Arnoldy, R. Hoffman, L. Peterson, and J. R. Winckler, Observations of low-energy solar cosmic rays from the flare of 22 August 1958, *J. Geophys. Research*, **64**, 1133-1147, 1959.
- Arnoldy, R. L., R. A. Hoffman, and J. R. Winckler, Observations of the Van Allen radiation regions during August and September 1959, part 1, *J. Geophys. Research*, **65**, 1361-1376, 1960.
- Freier, P. S., E. P. Ney, and J. R. Winckler, Balloon observation of solar cosmic rays on March 26, 1958, *J. Geophys. Research*, **64**, 685-688, 1959.
- National Bureau of Standards, *Solar-Geophysical Data, Part B, April and May, 1960*, Boulder, Colorado, 1960.
- Ney, E. P., J. R. Winckler, and P. S. Freier, Protons from the sun on May 12, 1959, *Phys. Rev. Letters*, **3**, 183-185, 1959.
- Winckler, J. R., Balloon study of high-altitude radiations during the International Geophysical Year, *J. Geophys. Research*, **65**, 1331-1359, 1960.
- Winckler, J. R., and P. D. Bhavsar, Low-energy solar cosmic rays and the geomagnetic storm of May 12, 1959, *J. Geophys. Research*, **65**, 2637-2655, 1960.

(Received July 1, 1960.)

# The Effect of the Initial Phase of a Magnetic Storm upon the Outer Van Allen Belt

J. M. MALVILLE

High Altitude Observatory, University of Colorado  
Boulder, Colorado

We present in this note the results of calculations of the longitudinal invariant of the motion of a charged particle trapped in a magnetic dipole field and in a compressed dipole field. The longitudinal adiabetic invariant [Chandrasekhar, 1958; Chew, Goldberger, and Low, 1955],

$$J = \oint V_{\parallel} ds$$

has already been used to advantage in describing other geophysical phenomena. Auroral isochasms have been satisfactorily described with the longitudinal invariant for the earth's main field [Vestine and Sibley, 1959]. Also, using the invariant [Dessler and Karplus, 1960], the gap in the outer Van Allen electron belt at  $1.9 \times 10^4$  km may be accounted for by the Capetown magnetic anomaly.

The component of velocity parallel to the magnetic field is given by

$$V_{\parallel} = \sqrt{(2\mu/m)(B_m - B)} \quad (1)$$

while the element of arc length of a dipole line of force is

$$ds = \sec^2 \phi_0 \sqrt{1 + 3 \sin^2 \phi} \cos \phi d\phi \quad (2)$$

$\mu$  is the magnetic moment of the particle,  $\phi_0$  the latitude of intersection of the line of force with the earth's surface, and  $B_m$  the magnetic intensity at the mirror height. The invariant may be expanded

$$J = \sqrt{\frac{2\mu}{m} B_m} \oint ds \left[ 1 - \frac{1}{2} \left( \frac{B}{B_m} \right) - \frac{1}{8} \left( \frac{B}{B_m} \right)^2 - \frac{1}{16} \left( \frac{B}{B_m} \right)^3 \right] \quad (3)$$

where  $B$  is given by

$$B = \frac{0.32[1 + 3 \sin^2 \phi]^{1/2}}{\sec^6 \phi_0 \cos^6 \phi}$$

The results of the integration of (3) for latitudes  $80^\circ$  to  $50^\circ$  and a mirror height of 1000 km are shown in Figure 1.

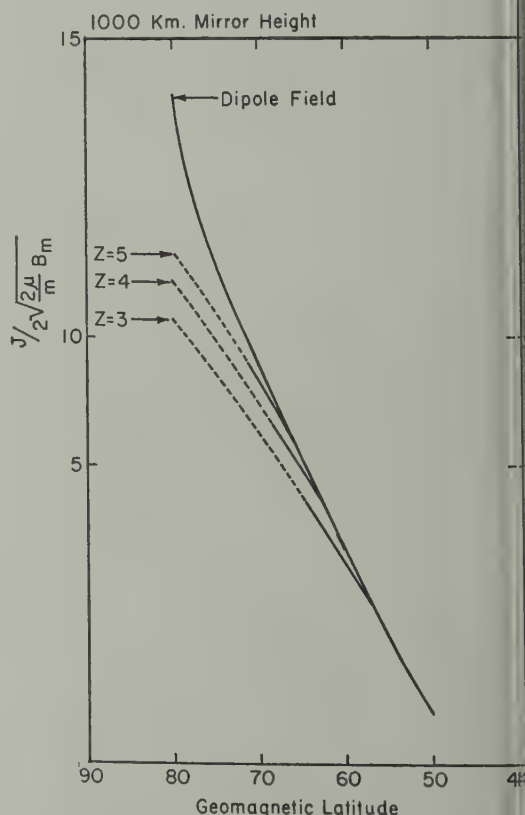


Fig. 1. Variation of the longitudinal invariant with  $\phi_0$ , the latitude of intersection of the line of force with the earth's surface, calculated for dipole field and for a field compressed by an image dipole at a distance  $2Z$  from the dipole axis.

In the Chapman-Ferraro theory of the initial phase of a magnetic storm the compressed day-side field is a sum of the earth's field plus that of an image dipole placed at a distance  $2Z$  from the geomagnetic axis,  $Z$  being the distance from the axis to the front of the charged plane. The latitude,  $\phi_1$ , at which the normal dipole lines of force strike the charged plane is given by

$$\cos \phi_1 = [Z \cos^2 \phi_0]^{1/3}$$

For the portion of the trapped particle's orbit parallel to the plane, the invariant is

$$J' = \sqrt{\frac{2\mu}{m}} B_m 2 \cdot \int_0^{\phi_1} \left[ 1 - \frac{B}{B_m} \right]^{1/2} Z \sec^2 \phi \, d\phi \quad (4)$$

In the cases we are considering ( $Z \geq 3$ ),

$$[1 - (B/B_m)]^{1/2} \sim 1$$

and we obtain

$$J' = 2 \tan \phi_1 \sqrt{(2\mu/m) B_m Z}$$

For  $\phi > \phi_1$  we are justified in using the invariant for the normal dipole field because of the  $1/r^3$  dependence of the field strength of the image dipole. To a good approximation the field inside the charged plane is that of the earth's dipole field alone. According to the Chapman-Ferraro theory, induction currents set up in the advancing charged sheet would produce a vanishing field at the surface of the sheet along the line of force specified by

$$\sin^2 \phi'_0 = 0.54/Z$$

Hence  $\phi'_0$  would represent the poleward limit of the trapped radiation belt. However, the finite conductivity of the sheet and the interplanetary medium may modify the situation so that the rate of diffusion of the field into the solar plasma may predominate in determining the extent of the outer Van Allen belt [Piddington, 1959]. For various values of  $Z$  we show in Figure 1 the invariant in the compressed dipole field; a broken line is used for latitudes  $\phi_0 > \phi'_1$ .

It is clear from Figure 1 that no latitude shift of the trapped radiation greater than  $2^\circ$  can be expected for even the strongest of magnetic

storms if one requires that  $\phi_0 < \phi'_1$ . This result is at variance with the suggestions of Rees and Reid [1959] that a latitude shift of  $10^\circ$  to  $15^\circ$  should be expected between the compressed dipole on the day side and the normal dipole on the night side. Such a latitude shift would be an attractive explanation for the diurnal variation of the latitude of auroral hydrogen emission observed by Reid and Rees [1960] and for the apparent spiraling zones of auroral, ionospheric, and magnetic activity in the polar cap [Malville, 1959]. Likewise, Vestine [1960] has suggested that a shift of the trapped radiation toward lower latitudes on the night side would lower the mirror heights and produce a maximum dumping of the trapped particles near

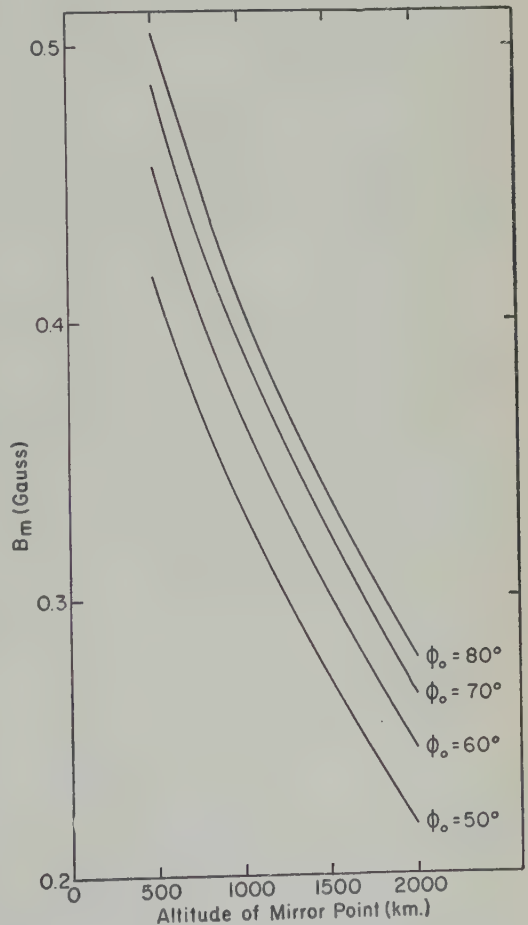


Fig. 2. Magnetic field strength in a dipole field at the elevation of the mirror point.



geomagnetic midnight. The occurrence of maximum auroral activity at magnetic midnight appears not to be under strong terrestrial longitude control, and using the dipole approximation to the earth's main field we may obtain an estimate of the 'dumping' effect. In Figure 2 we show the variation of  $B_m$  with altitude and magnetic latitude for a dipole field. Even were a latitude shift of  $10^\circ$  to  $15^\circ$  common, it would indeed have negligible effect in lowering  $B_m$  into auroral elevations in the atmosphere; for instance, a particle originally mirroring at 1000 km at  $70^\circ$  latitude would drop only to 850 km if shifted to the  $60^\circ$  line of force.

It is likely, however, that at least during the main phase of a magnetic storm the field is distended either through the current of trapped particles [Dessler and Parker, 1959; Akasofu, 1960] or through the creation of a geomagnetic tail [Piddington, 1960]. An extension of the lines of force would increase  $J$  for a given  $\phi_0$  and would increase the equatorward shift of the trapped articles. If the lines of force were extended in length by a factor of 4 beyond 4 earth radii, a particle originally trapped at the day side (with  $Z = 5$ ) along a line of force  $\phi_0 = 71^\circ$  would shift to  $63^\circ$ .

Especially with a rotating earth, it is not likely that such lengthened lines of force would maintain sufficient identity or continuity to be able to trap particles. In view of the lack of an adequate theory for or description of the needed large extension of the lines of force, we cannot, at present, associate the observed diurnal latitude shifts of various geophysical phenomena with a shift of the trapped radiation belt.

This work was carried out while the author held a fellowship from the Institute for Solar Terrestrial Research of the High Altitude Observatory.

## REFERENCES

- Akasofu, Syun-Ichi, The ring current and the outer atmosphere, *J. Geophys. Research*, **65**, 535-544, 1960.
- Chandrasekhar, S., *The Plasma in a Magnetic Field*, edited by R. K. M. Landshoff, Stanford University Press, 1958.
- Chew, G. F., M. L. Goldberger, and F. E. Low, *Los Alamos Rept. LA-2055 T-767*, 1955.
- Dessler, A. J., and R. Karplus, Some properties of the Van Allen radiation, *Phys. Rev. Letters*, **271-273**, 1960.
- Dessler, A. J., and E. N. Parker, Hydromagnetic theory of geomagnetic storms, *J. Geophys. Research*, **64**, 2239-2252, 1959.
- Malville, J. M., Antarctic auroral observations at Ellsworth Station, 1957, *J. Geophys. Research*, **64** (10), 1389-1393, 1959.
- Piddington, J. H., The transmission of geomagnetic disturbances through the atmosphere and interplanetary space, *Geophys. J.*, **2**, 173-188, 1959.
- Piddington, J. H., Geomagnetic storm theory, *Geophys. Research*, **65**, 93-106, 1960.
- Rees, M. H., and G. C. Reid, The aurora, the radiation belt and the solar wind: A unifying hypothesis, *Nature*, **184**, 539-540, 1959.
- Reid, G. C., and M. H. Rees, Spectroscopic development of the aurora, *Planetary Space Science*, in press, 1960.
- Vestine, E. H., Polar auroral, geomagnetic, and ionospheric disturbances, *J. Geophys. Research*, **65**, 360-361, 1960.
- Vestine, E. H., and W. L. Sibley, Remarks on auroral isochasms, *J. Geophys. Research*, **64**, 1338-1339, 1959.

(Received May 27, 1960.)

## Increase of the Nucleonic Intensity on May 4, 1960

J. H. TRAINOR, M. A. SHEA, AND J. A. LOCKWOOD<sup>1</sup>

*Department of Physics, University of New Hampshire  
Durham, New Hampshire*

An increase in the cosmic-ray intensity at the earth was observed on May 4, 1960, ~16 minutes after a class 2 solar flare at ~1015 UT in region H42 on the west limb, N10° heliographic latitude (Preliminary Report of Solar Activity, High Altitude Observatory, Boulder, Colo.). This intensity increase was very rapid and was recorded only at certain stations within well-defined regions on the earth (K. G. McCracken and R. Palmeira, private communication; H. Carmichael and J. F. Steljes, preliminary report). We report here the observations on the increase in the nucleonic intensity at Mt. Washington (elevation, 1909 m) and Durham (sea level), and some tentative conclusions regarding the rigidity spectrum of the flare-produced particles.

The nucleonic intensity at Mt. Washington for the period of the major increase is plotted in Figure 1. Also given are the times of the flare and the 27.6-Mc/s riometer events at Kiruna, Sweden (J. Ortner, private communication) and

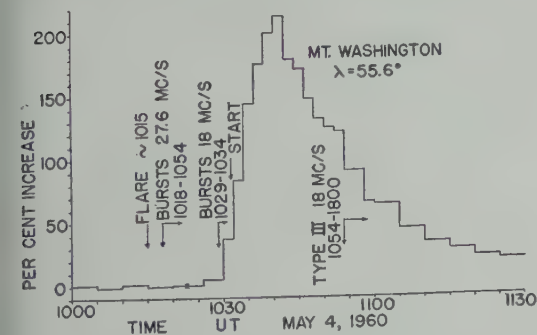


Fig. 1. The increase in nucleonic intensity at Mt. Washington. Significant events recorded by 18-Mc/s riometers are also given.

<sup>1</sup> On leave at the Royal Institute of Technology, Stockholm, Sweden.

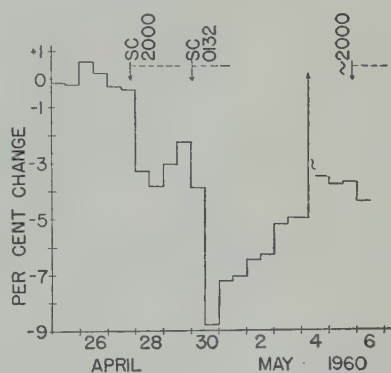


Fig. 2. The 12-hour average neutron intensity as a per cent change from the average of May 25-27, 1960. The onset and duration of geomagnetic events are noted.

the 18-Mc/s events at Rensselaer Polytechnic Institute, Troy, New York (R. E. Falconer, private communication). The cosmic-ray increase started at  $1031.5 \pm 1.0$  UT and reached a maximum increase of 214 per cent between 1040 and 1042, falling off to 18 per cent between 1125 and 1130. In the period 1830 to 1930, the intensity was still ~2.4 per cent above the pre-increase level. At Durham the onset time was  $1032 \pm 2$  or  $-1$  UT, and the magnitude as determined from a calibrated counting rate meter was  $140 \pm 10$  per cent, reaching a maximum at ~1040. From the calculations by *Frior* [1954] we estimate that both these stations were in the 0400 zone.

The nucleonic intensity on May 4 was about 20 per cent below that in 1954. As is shown in Figure 2, the intensity was recovering from a series of Forbush events, none of which appears to be associated with the region that produced the cosmic-ray increase. The sudden-onset geomagnetic storm that began ~0200 on May 5 can be attributed to the H42 flare, how-

ever (High Altitude Observatory, Boulder, Colo.).

Since Mt. Washington and Durham have essentially the same vertical cutoff rigidity, we can use the altitude dependence of the increase to determine the differential rigidity spectrum of the solar particles. If we make the assumptions that the rigidity spectrum is given by  $K/P^n$  and that there are only protons present, then we can use the specific yield functions for a nucleonic detector deduced by *Webber and Quenby* [1959] to determine  $n$ . The differential counting rate,  $dN/dP$ , for 820 g cm<sup>-2</sup> was determined, and, from this, the proton specific yield function was computed for Mt. Washington. Just before the increase, the primary rigidity spectrum was different from that in 1954, and this was adjusted using a modulation factor of  $1.5/P$ . It was then possible to determine the ratio of Mt. Washington to Durham for various values of  $n$ . The observed ratio of the increases at Mt. Washington to Durham was 1.43–1.65 at the maximum and 1.6 for the period 1030–1130 UT, corresponding to a rigidity spectrum with  $n$  equal to 5.

We can make a further comparison with the intensity increase of ~5.5 per cent recorded by the Massachusetts Institute of Technology meson detector (K. G. McCracken and R. Palmeira, private communication) using *Webber and Quenby's* [1959] data for the ionizing component, since MIT, Durham, and Mt. Washington are located close together. The ratios of the intensity increases in the nucleonic component at

Mt. Washington and Durham to the MIT meson telescope were 39 and ~25, respectively, consistent with an exponent of 5.

Applying the same method to data for the February 23, 1956, flare with the modified cutoff rigidities of *Quenby and Webber* [1959], we found  $n$  to be 6–7, agreeing with the measurements of *Meyer, Parker, and Simpson* [1956]. We believe that these estimates indicate that the differential rigidity spectrum for the flare produced particles was  $K/P^6$ , not as steep in the 1956 event.

*Acknowledgments.* The work for this letter was supported by the Geophysical Research Directorate of the Air Force Cambridge Research Center and the U. S. National Committee for the IGY.

We are indebted to the High Altitude Observatory and the National Bureau of Standards, Boulder, Colorado, for the Preliminary Report of Solar Activity.

#### REFERENCES

- Frior, J., Cosmic radiation intensity-time variations and their origin, IV, Increases associated with solar flares, *Phys. Rev.*, **94**, 1017–1028, 1953.
- Meyer, P., E. N. Parker, and J. A. Simpson, Solar cosmic rays of February 1956 and their propagation through interplanetary space, *Phys. Rev.*, **104**, 768–783, 1956.
- Quenby, J. J., and W. R. Webber, Cosmic ray cut-off rigidities and the earth's magnetic field, *Phil. Mag.*, **4** (37), 90–113, 1959.
- Webber, W. R., and J. J. Quenby, On the derivation of cosmic ray specific yield functions, *Phil. Mag.*, **4** (41), 654–664, 1959.

(Received June 4, 1960.)

## A Search for Geomagnetic Singular Days

T. POHRTE

*Colorado School of Mines, Golden, Colorado*

C. WARWICK

*National Bureau of Standards, Boulder Laboratories  
Boulder, Colorado*

AND

N. MACDONALD

*High Altitude Observatory, University of Colorado  
Boulder, Colorado*

I. A study by *Shapiro and Ward* [1960] shows that the average value of the daily geomagnetic character figure  $C_i$  for a given calendar date shows no departures from a smoothed value beyond the deviations expected by chance. Furthermore, deviations of the daily mean value in one subsample show little correlation with the deviations in another subsample. These authors conclude that, apart from the well-known seasonal variation, there are no significant deviations in geomagnetic activity for a given calendar date, or no 'singular days' in geomagnetic activity.

We have carried out a series of similar studies that serve to confirm and strengthen this conclusion. Various indices of geomagnetic disturbances were used. In particular, we looked for a possible difference between years of high solar activity and years of lower activity. Because the character of geomagnetic disturbance changes with phase of the solar-activity cycle, significant effects can be obscured in a statistical treatment that combines the SC type of disturbance, prevalent at time of high solar activity, with the recurrent storms that characterize years of low solar activity.

All studies showed deviations of the type illustrated and discussed by Shapiro and Ward. We will not discuss here the probability of occurrence of such deviations, but will show that the dates of large deviations do not correspond from one study to another.

II. *Description of indices.* In each study, each calendar date is characterized by an index that

measures geomagnetic activity on that date over a number of years. These various indices are described below.

A. Studies of all years combined:

1. The number of days with storm commencement, 1890–1957. A day for which  $C_i \geq 1.0$  and is greater by at least 0.5 than  $C_i$  for the previous day was defined as the day of commencement of disturbance.

2. Number of days of large  $A_p$ , 1940–1958. Days for which  $A_p \geq 30$ .

3. Mean value of  $A_p$ , for each day, 1940–1958.

B. Studies of recurrent storms, and years of low solar activity:

1. Number of days with commencement of disturbance, 1890–1957. A disturbance commencement is defined by  $C_i$  and  $\Delta C_i$  as in A1. The day must also be a member of a series of at least 3 such days following at intervals of  $27 \pm 3$  days.

2. Number of days with storm beginning, during years of low solar activity, 1874–1954. Days were selected from lists of beginning dates of 'great' and 'small' magnetic storms as observed at Greenwich Observatory [*Royal Greenwich Observatory*, 1955]. For each ascending and descending branch of the sunspot cycle, the time was found when the smoothed monthly mean sunspot number passed the mean of the values for the maximum and minimum of the cycle. The periods between these times were defined as years of high activity (around maximum) and years of low activity (around minimum).



TABLE 1

| Dates of High Index   |                        |                       | Dates of Low Index    |                       |                    |          |
|-----------------------|------------------------|-----------------------|-----------------------|-----------------------|--------------------|----------|
| A1 ( $C_i$ )          | A2 ( $A_p > 30$ )      | A3 ( $\bar{A}_p$ )    | A1 ( $C_i$ )          | A2 ( $A_p > 30$ )     | A3 ( $\bar{A}_p$ ) |          |
| All years             | March 01               | Jan. 18               | May 16                | April 19              | Feb. 09            | Mar. 19  |
|                       | March 21               | June 25               | June 25               | June 11               | Mar. 17            | Apr. 01  |
|                       | Sept. 03               | June 29               | July 01               | July 07               | Apr. 14            | Sept. 10 |
|                       | Sept. 30               | June 30               | Sept. 03              | July 19               | Oct. 06            | Oct. 00  |
|                       | Oct. 18                | Sept. 03              | Oct. 28               | Oct. 30               | Nov. 05            | Oct. 14  |
|                       |                        |                       | Nov. 30               |                       |                    |          |
| Years of low activity | B1 ( $C_i$ )           | B2 (Greenwich storms) | B1 ( $C_i$ )          | B2 (Greenwich storms) |                    |          |
|                       | March 01               | Jan. 19               | Feb. 18               | Mar. 07               |                    |          |
|                       | March 21               | Jan. 24               | Feb. 29               | Apr. 04               |                    |          |
|                       | May 04                 | Apr. 06               | May 19                | Sept. 06              |                    |          |
|                       | Sept. 03               | June 29               | June 15               | Sept. 12              |                    |          |
|                       | Sept. 30               | Oct. 23               | June 24               | Oct. 09               |                    |          |
|                       |                        | Oct. 25               | July 18               |                       |                    |          |
|                       |                        | Oct. 28               | Aug. 24               |                       |                    |          |
|                       |                        |                       | Sept. 02              |                       |                    |          |
|                       |                        |                       | Sept. 08              |                       |                    |          |
|                       |                        |                       | Sept. 29              |                       |                    |          |
|                       |                        |                       | Oct. 17               |                       |                    |          |
|                       |                        |                       | Nov. 07               |                       |                    |          |
|                       |                        |                       | Nov. 30               |                       |                    |          |
|                       | Years of high activity | C1 ( $C_i$ )          | C2 (Greenwich storms) |                       |                    |          |
| March 24              |                        | Jan. 25               |                       |                       |                    |          |
| May 26                |                        | May 05                |                       |                       |                    |          |
| Oct. 10               |                        | Aug. 20               |                       |                       |                    |          |
| Oct. 27               |                        | Aug. 22               |                       |                       |                    |          |
|                       | Nov. 01                | Oct. 13               |                       |                       |                    |          |

Monthly mean values and maximum and minimum values of sunspot number were obtained from the study of *Chernosky and Hagan* [1958].

C. Studies of nonrecurrent storms and years of high solar activity:

1. Number of disturbance commencements that do not fall in a recurrent series (members of A1 but not B1), 1890–1957.

2. Number of beginnings of large storms described by the Greenwich Observatory during years of high solar activity, 1874–1954. Storms were selected from the list of ‘great’ and ‘small’ storms for which  $\Delta D \geq 50'$ ,  $\Delta H$  or  $\Delta Z \geq 250\gamma$ .

III. Selection of days of large deviation. In each study, the daily index was plotted against the date. A smooth curve representing the semi-annual variation was drawn through these points. The standard deviation was estimated at each

point of the smooth curve, and points that deviated the most from the smooth curve, in terms of the standard deviation, were selected.

In both studies in section C, there were large number of dates with no disturbance (index = 0), and so no dates with especially large negative deviations are listed for this group. Each heading in Table 1 consists of the section described in part II, followed, in parentheses, the index used in selecting the dates. In the table are given dates of large deviations. In general, 5 high and 5 low days were selected from each study, but in some cases more or fewer were considered outstanding.

The table shows that there is no one day that consistently shows large deviation in geomagnetic activity. We conclude that there are singular days for geomagnetic disturbance.

We are grateful to Dr. Walter Orr Roberts who suggested this study.

This work was supported by the Air Force Cambridge Research Center, the High Altitude

Observatory Institute for Solar-Terrestrial Research, and the National Bureau of Standards.

## REFERENCES

Chernosky, E., and M. Hagan, The Zurich sunspot number and its variations for 1700-1957, *J. Geophys. Research.* 63, 775-788, 1958.

Royal Greenwich Observatory, *Sunspot and Geomagnetic-Storm Data*, Her Majesty's Stationery Office, London, 1955.

Shapiro, R., and F. W. Ward, Jr., Daily normals of the geomagnetic character figure  $C_4$ , *J. Geophys. Research*, 65, 115-117, 1960.

(Received June 4, 1960; revised July 5, 1960.)

Measurements of the Geomagnetic Field near Capetown<sup>1</sup>

J. R. HEIRTZLER AND J. HIRSHMAN

*Lamont Geological Observatory  
(Columbia University)  
Palisades, New York*

It has recently been suggested [Dessler, 1959; Dessler and Karplus, 1960] that the anomalously weak geomagnetic field near Capetown, South Africa, is responsible for an observed particle flux gap in the electron belt of the Van Allen radiation zones. This anomaly may also

allow neutron decay electrons to be supplied to the outer zone. Mirror altitudes of the trapped radiation are quite sensitive to the intensity of the total magnetic field. Recent geomagnetic measurements covering a large geographic area are scarce for the Capetown anomaly, and we would like to make a preliminary report of magnetic measurements made off the southwest coast of Africa. We know of no new magnetic data from

<sup>1</sup> Lamont Geological Observatory Contribution No. 425.

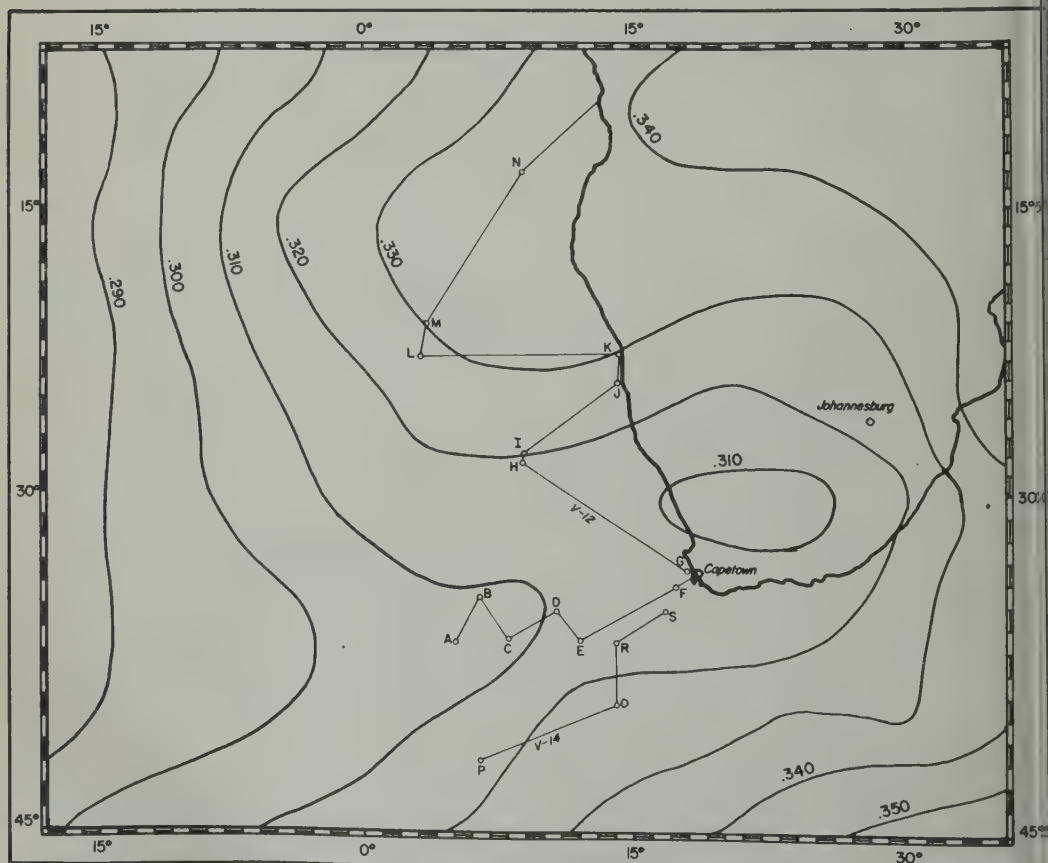


Fig. 1. Tracks of the 1957 (V-12) and 1958 (V-14) cruises of the R.V. *Vema* southwest of Africa. The total magnetic field contours are from the 1955 U. S. Hydrographic Chart of total intensity.

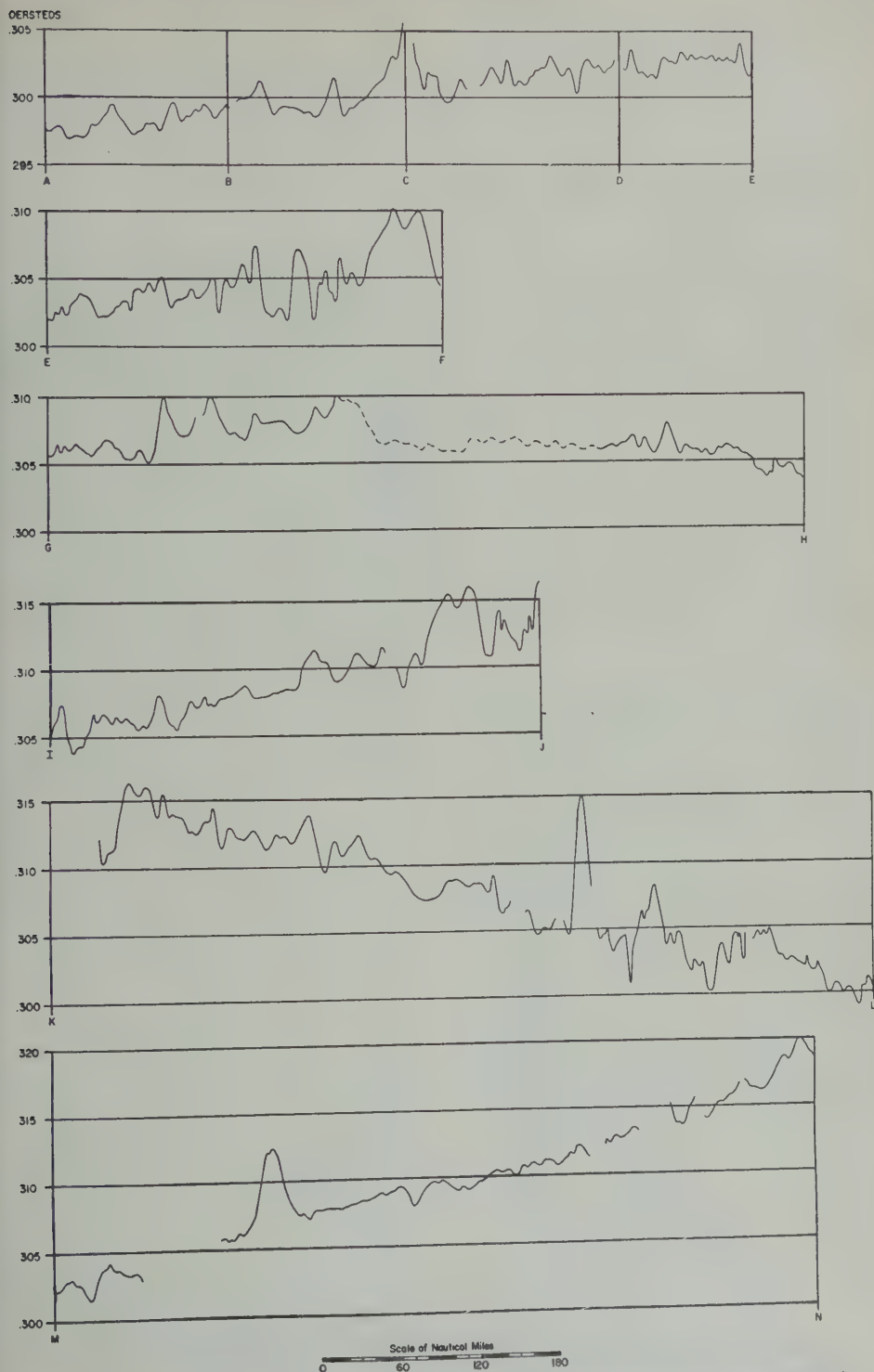


Fig. 2. Total magnetic field intensity as a function of distance between the points A through N shown in Figure 1.



this oceanic area published in recent years. The data reported here were acquired on the cruises of the R.V. *Vema* during 1957 and 1958. Some discrepancy is found with the most recent U. S. chart of the total field intensity.

The tracks of the *Vema*-12 cruise (1957) and the *Vema*-14 cruise (1958) are shown in Figure 1. Small side excursions have been omitted from the illustration. Total field contours taken from the U. S. Hydrographic Chart No. 1703 of the total field intensity for 1955 are displayed on the same figure. The annual field changes indicated on the 1955 chart are not shown.

The total magnetic field intensity was measured at sea level with a towed flux-gate magnetometer [Miller and Ewing, 1956]. Measurements were taken continuously except for brief periods when the instrument was being serviced or other geophysical work interfered. The total field as a function of distance along each leg of the tracks is given in Figures 2 and 3. The results reported here have an absolute error of less than  $100 \gamma$  ( $1 \gamma = 10^{-5}$  oersted) and a relative error during each cruise of less than  $10 \gamma$ . Because of navigational problems the dashed segment of the line in Figure 2 is less accurate than other values but is probably within this error.

The localized fluctuations are believed to be of geological origin. The part of the *Vema*-12 cruise was completed during the period April 19 to May 13, 1957. The  $A_p$  index did not exceed a value of 27 ( $54 \gamma$ ) for any day during this period, except on April 19 when it reached a value of 60

( $120 \gamma$ ). The *Vema*-14 track was covered during March 27 to April 6, 1958, and during the second period the  $A_p$  index did not exceed 27 ( $54 \gamma$ ). We do not feel that any of the fluctuations exceeding  $150 \gamma$  are due to short time variations of the field.

We hesitate to draw 1000- $\gamma$  contours from the limited data presented, but it is clear that the 1955 contours, corrected for the annual change indicated on the 1955 chart, are not in accord with the values reported here. In some areas the 0.310-oersted contour for 1955 should be moved to the east by  $17^\circ$ . Our values appear to coincide more closely with the U. S. S. R. isomagnetic maps for 1955 than with the U. S. maps. However, extensive information from the Soviet Union was not available.

For past epochs charts of annual change of total intensity (isoporic charts) have foci which generally show a westward movement. A positive isoporic focus situated in the western part of the South Atlantic Ocean has shown a more irregular behavior than most [Fleming, 1948]. Accordingly the geomagnetic field in the South Atlantic appears to warrant careful and frequent surveying. Project Magnet, which is being conducted by the U. S. Hydrographic Office, is expected to include an aeromagnetic survey of this area.

Magnetic data for a 1959-1960 *Vema* cruise in the Capetown area have been acquired but not yet processed.

We would like to acknowledge the assistance of Carolyn Peppin. The work was done under contract with the Office of Naval Research.

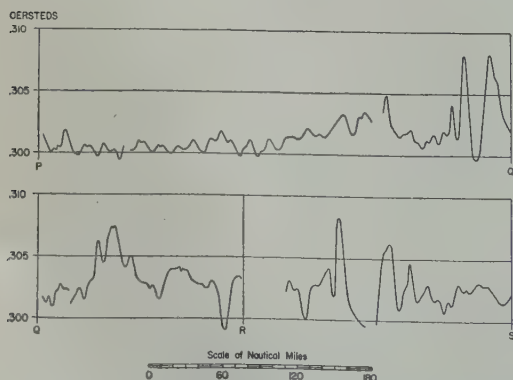


Fig. 3. Total magnetic field intensity as a function of distance between the points P through S shown in Figure 1.

## REFERENCES

- Dessler, A. J., Effect of magnetic anomaly on particle radiation trapped in geomagnetic field, *J. Geophys. Research*, **64**, 713-715, 1959.
- Dessler, A. J., and Robert Karplus, Some properties of the Van Allen radiation, *Phys. Rev. Lett.*, **4**, 271-274, 1960.
- Fleming, J. A., *Trans. Assoc. Terr. Mag. & Elect. Inter. un. Geod. Geophys.*, Oslo, p. 37, 1948 (as cited by S. K. Runcorn in *Encyclopedia of Physics*, **47**, p. 512, Springer-Verlag, 1956).
- Miller, E. T., and Maurice Ewing, Geomagnetic measurements in the Gulf of Mexico and in the vicinity of Caryn Peak, *Geophysics*, **21**, 400-432, 1956.

(Received July 5, 1960.)

## A Proposed X-Ray Telescope for the 1- to 100-A Region

ALBERT V. BAEZ

*Smithsonian Astrophysical Observatory  
Cambridge, Massachusetts*

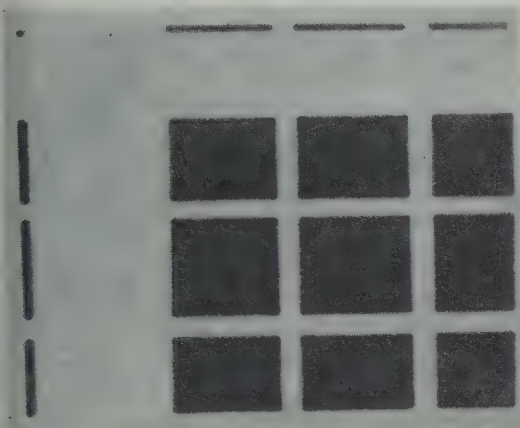


Fig. 1. The image of a small hole illuminated by 2537-A radiation made by the plane-mirror-multiple-crossed-mirror system. The large area with thin white lines through it was produced by the direct beam. Above this is the exposed region produced by the rays that were reflected only by the first set of mirrors. To the left of the direct beam is the region produced by rays that were reflected only by the second set of mirrors. In the upper left-hand corner is the image of the source. It was formed by light reflected from both sets of mirrors.

The recent pinhole radiograph of the sun made by *Friedman* [1960] clearly points up the need for image-forming devices to focus on solar and other X-ray sources in space. In the wavelength region from about 100 to 1000 Å the zone plate described by *Baez* [1960] should work well for intense sources. But the zone plate is a highly chromatic device, and therefore it is limited to the study of sources that emit line spectra or radiation that can be monochromatized. There is need for achromatic image-forming devices that are faster than the zone plate and can operate in the shorter-wavelength region from about 1 to 100 Å. Two such devices are described below. A crude prototype of the first has been built; it was used to make

the pictures shown in Figures 1 and 2. The second device is a contemplated refinement of the first, for which components are being designed and constructed.

The first telescope consists of two thin, plane glass mirrors, each approximately 1 inch by 3 inches, disposed like the shades of a Venetian blind, followed by an identical array at right angles to the first. The effect of the first set alone would be to take X rays from a point source at infinity and focus them on a line whose width is determined, to a first approximation, by the product of the mirror width and the sine of the grazing angle of incidence. The



Fig. 2. To show that a multiple-crossed-mirror system is capable of producing an image of an extended object, an opening in the shape of a question mark was illuminated, by transmission, with white light. Its image appears in the upper left. The other regions represent the direct beam and the beams that experienced reflection at only one of the two mirror arrays, as explained in Figure 1.

long edges of the mirrors are parallel, but the normals to the reflecting surfaces make slightly different angles with the incident beam so that all reflected beams coincide in a chosen focal plane normal to the reflected beams. The effect of the second set of mirrors is to take the rays reflected by each of the first mirrors and reflect them into a small rectangle. This rectangle represents the approximate image of the point source at infinity. Both its height and its width are determined by the mirror widths and the grazing angles of incidence. With X rays, the mirror widths can be made very small before diffraction effects start widening the beam. Therefore, good resolution is possible.

Figure 1 shows the image of a small hole illuminated by 2537-A radiation made by the multiple-crossed-mirror system. The large area with thin white lines through it was produced by the direct beam, that is, by that portion of the incident beam that missed both the mirror systems. Above this direct beam can be seen the area produced by the rays that were reflected by the first set of mirrors but missed the second set. To the left of the direct beam is an area produced by the rays that hit the second mirror system but missed the first. In the upper left-hand corner is the spot formed by the light that struck both sets of mirrors; it represents the image of a small source.

To show that the system is capable of producing an image of an extended object, we illuminated an opening in the shape of a question mark. Its image, formed by a multiple-crossed-mirror system with four mirrors in each array and using white light, is shown in Figure 2. The rays that experienced only single reflections can be identified by referring to the remarks made about Figure 1.

With this multiple-crossed-mirror system the image of a distant point source resembles that produced by a rectangular pinhole, since each crossing of two mirrors at right angles to each other presents effectively a small rectangular 'window' for the radiation. From the point of view of total flux gathered, a system with  $n$  mirrors in each array is equivalent to  $n^2$  pinholes, each pouring all its energy into the same spot.

The advantages of the system include simplicity of construction, approximately true point-to-point image-forming characteristics, absence of chromatic aberration, and speed (large flux-gathering capabilities).

The second telescope, a contemplated improvement of the first, is to be made by curving the mirrors so that each one effectively becomes a section of a right-circular cylinder. The axes of all the cylinders would be parallel to one another and perpendicular to the incoming beam. Curving the mirrors will greatly improve the resolving power of the system and the intensification of illumination at the image. At grazing incidence, spherical surfaces may be used instead of cylindrical ones with no practical loss in resolution, as has been demonstrated by Kirkpatrick and Baez [1948] and Kirkpatrick and Pattie [1957] with reflection X-ray microscopes.

Spherical mirrors are easier to grind than cylindrical ones and are more stable than cylindrical mirrors made by bending flat plates in the manner of Ehrenberg [1947]. Experience with crossed-mirror systems in X-ray microscopy has shown that they can be effective even at wavelengths as short as 1.5 Å. Many of the restrictions on these systems are eased at longer wavelengths; therefore, the useful range of multiple-crossed-mirror systems is likely to be 1 to 100 Å or beyond. A detailed summary of theoretical and experimental results will appear in the *Journal of the Optical Society of America*.

#### REFERENCES

- Baez, A. V., A self-supporting metal Fresnel zone plate to focus extreme ultra-violet and soft X-rays, *Nature*, **186**, 958, 1960.
- Ehrenberg, W., X-ray optics, *Nature*, **160**, 332, 1947.
- Friedman, H., Survey of observations of soft ultraviolet and X-rays, Symposium on Solid Emissions and the Interplanetary Medium, *Nat. Acad. Sciences, Am. Geophys. Union*, April 22, 1960.
- Kirkpatrick, P., and A. V. Baez, Formation of optical images by X-rays, *J. Opt. Soc. Am.*, **33**, 766, 1948.
- Kirkpatrick, P., and H. H. Pattie, Jr., X-ray microscopy, in *Handbuch der Physik*, edited by S. Flügge (*Encyclopedia of Physics*), Springer-Verlag, Berlin, 1957.

(Received June 9, 1960.)



**Author's Reply to the Discussion by M. A. Benson  
of the Paper, 'A New Approach to Peak Flow Estimation'**

R. RANGARAJAN

*Central Water and Power Commission  
New Delhi, India*

I am grateful to *Benson* [1960] for his keen interest in the new approach [*Rangarajan*, 1960] to an old problem and also for a few doubts raised by him which are explained in the following paragraphs.

This paper was the sequel to my attempt to study the effects of intercorrelations of the daily flow data on the flood estimates made by Gumbel's extreme value distribution. On account of the complicated mathematics involved in adjusting the Gumbel distribution itself for the intercorrelation effects, I tried to approach the problem from a new angle, using daily flow records as explained in the paper. I myself was puzzled by the small amount of change in the flood estimates effected by the introduction of the intercorrelations as compared with those obtained when no correlation was assumed. The analysis showed that, even if the intercorrelations were higher than 0.85, the effect would not be appreciable. But it was far more revealing and interesting to find quite ample divergence between Gumbel's estimation and that by the new approach when  $\rho = 0$ . Benson attributes the divergence either to the inherent difference in what the two methods represent, rather than to the defects in Gumbel's formula, or to the analysis being restricted to only 3 months. The latter explanation is not valid, since the exceedance probabilities obtained for the data earlier than July and also after September will be too small to increase the compounded probability materially. Regarding the first explanation that there are inherent differences in the two procedures, it may be noted that the distribution functions (3) and (4) are the exact distribution of the extreme value when  $\rho = 0$ , whereas the Gumbel formula is only a limiting form of the actual extreme value distribution. Hence there is every reason to expect the new approach to give more precise estimates than Gumbel's. Really, a part

of the discrepancy arises because of the approximation involved in Gumbel's expression, whereas the nonstationarity of the means of log discharges on various dates accounts for the remaining discrepancy.

Benson has questioned the oversimplifying assumptions made in the derivation of the various probabilities. Though three broad assumptions were postulated for obtaining a simplified formula, some of the conditions can be relaxed to a fair degree to accommodate more general problems, as illustrated in the Yamuna example.

The use of an average standard deviation for all different dates is justified because the standard deviation plotted in Figure 3(b) is quite random and has no trend movements as in the case of means in Figure 3(a). Further, the return period of 200,000-ft<sup>3</sup>/sec floods was worked out by compounding the probability for the different dates when the flow value on each of the dates is distributed according to normal law with the mean and standard deviation actually obtained for that date from the available observations. The return periods were found to differ very little from those obtained with the group means for the different dates and one average standard deviation for the entire period of 3 months.

The small increasing trend observed in the correlation does not have any physical explanation. Also, the use of different correlations for different periods, while making the calculations very laborious, will not give results in any way different from those obtained using average correlation inasmuch as there is no similarity in the long-term movements of the means and the correlation coefficients; i.e., the same value of group mean is pitted once against a small correlation in the first half of the period and at another time against higher correlation at the later period, and this compensating effect will nullify



the differences obtained by using an average correlation instead of varying correlations.

The condition (2) in the paper implies the stationarity of the joint distribution, including the parameters, which is not true for cases like the Yamuna, where the means on different dates are not the same. Hence, while applying the method, it was assumed that the distribution is the same but for changing the location parameters (means). This assumption, I believe, is quite justifiable.

Thus I believe that the new approach is help-

ful in throwing light on the seasonal and inter-correlation effects and in exploring the possibilities of improving the estimates obtained by classical methods such as Gumbel's.

#### REFERENCES

- Benson, M. A., Discussion of paper by R. Rangarajan, 'A new approach to peak flow estimation', *J. Geophys. Research*, 65, 2557, 1960.  
Rangarajan, R., A new approach to peak flow estimation, *J. Geophys. Research*, 65, 643-651, 1960.

(Received July 1, 1950.)

# Electrical Resistivity of Frozen Earth<sup>1</sup>

T. HATHERTON

*Geophysics Division*

*Department of Scientific and Industrial Research*

*Wellington, New Zealand*

Cook [1960] has described laboratory experiments on the RF electrical properties of synthetic salty ice and frozen earth and has commented on the scarcity of available data concerning the electrical properties of natural sea ice and permafrost with which to compare his results. During IGY the d-c resistivities of the ground upon which Scott Base (latitude 77° 51'S) and Hallett Station (latitude 72° 19'S) are constructed were measured, and this note presents the d-c resistivities of the natural frozen ground and compares them with the RF resistivities of the synthetic soils used by Cook.

Pram Point, Ross Island, on which Scott Base is built, consists of basaltic debris, with a good deal of clay material, overlying the basalt flows of the Hut Point Peninsula. Except for the top 5 to 10 cm, which thaws for a very short period during the summer months, the ground is permanently frozen. Hallett Station lies on the gravel spit attached to Cape Hallett, and in at least the upper 2 meters of this soil there is some penguin guano. Resistivities measured at these bases during January 1958 by the standard d-c method (commutated) using the Wenner electrode configuration are shown in Figure 1. At Scott Base two expanding electrode profiles were measured in different areas on Pram Point. At Hallett Station only one locality was occupied.

At Scott Base resistivities near the surface are about 100 ohm-meters in one locality and about 300 ohm-meters in the other. Apparent resistivities at depth are 1500 to 3000 ohm-meters. At the time of measurement the ground temperature at 15 cm was -5°C; the temperature of the ground at depths below 10 meters can be taken to be about -21°C, the mean annual temperature. At Hallett Station the resistivity

near the surface was only about 30 ohm-meters, and at the greatest electrode separation ( $a = 30.5$  meters) the resistivity was about 400 ohm-meters.

At the time of the measurement it was thought that these d-c resistivities were surprisingly low for frozen ground; they are of the same order as the resistivities of surface waters and less than an order of magnitude greater than many ground waters [Birch, 1942]. However, the salt concentration in the subsurface ice is unknown.

The RF resistivities at 100 Mc/s of the various synthetic permafrosts measured by Cook ranged from 68 ohm-meters for high-Fe clay to 1200 ohm-meters for wet coarse sand containing tap water with only 0.2 per cent  $\text{Ca}(\text{HCO}_3)_2$ . These values are similar to the apparent d-c resistivities (30 to 3000 ohm-meters) of natural frozen ground reported here. Though Cook uses a lower temperature for his 'frozen ground'

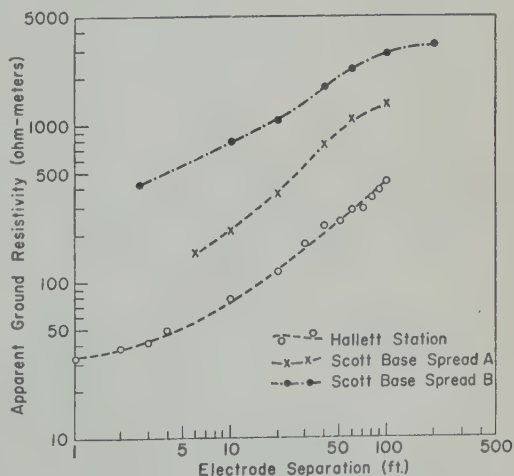


Fig. 1. Apparent electrical resistivities of the ground at Scott Base and Hallett Station, January 1958.

<sup>1</sup> Contribution No. 989, Division of Geological Sciences, California Institute of Technology.

( $-35^{\circ}\text{C}$ ) than existed in the frozen ground at Scott Base and Hallett Station, the resistivities of soils usually decrease with increasing frequency of the applied current [Birch, 1942, p. 316], so that these two factors approximately oppose each other and the two resistivity ranges are probably comparable. Cook also considers that the RF resistivities of his synthetic permafrost samples are much lower than he expected, though it is to be noted that his tap water has

an RF resistivity of 25,000 ohm-meters when frozen.

## REFERENCES

- Birch, F. (Ed)., *Handbook of Physical Constants*, Geol. Soc. Am. Spec. Paper 36, 1942.  
Cook, J. C., RF electrical properties of salty ice and frozen earth, *J. Geophys. Research*, 65, 1763-1771, 1960.

(Received June 24, 1960.)



# AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Avenue, N.W., Washington 5, D. C.

*Established by the National Research Council in 1919 for the development of the science of geophysics through scientific publication and the advancement of professional ideals.*

## QUALIFICATIONS FOR MEMBERSHIP

The membership of the AGU shall consist of Members, Associate Members, Student Members, and Corporation Members.

Those eligible as candidates for election to the grade of MEMBER shall be:

**MEMBER** (a) Persons who have made an active contribution to geophysical research through observation, publication, teaching, or administration. Definite evidence should be presented to the Membership Committee. "Publication" may include books, articles, unpublished manuscripts, inventions, or development of geophysical instruments.

(b) Persons who have made active practical application of geophysical research. It should be shown that the nominee's work has not been purely routine, but that it has tended to create new knowledge of, or to broaden or strengthen the application of, geophysical research. In general, the minimum qualifications for membership will be not less than three years of professional experience in some phase of geophysics.

*(Continued on next page)*

Cut along this line

## APPLICATION FOR MEMBERSHIP

Please refer to qualifications on reverse side and designate below type of membership desired:

Member (\$10) ☐

Associate (\$10) ☐

Student (\$4.50) ☐  
(1960)

Application forms for Corporation Membership are available upon request.

|                   |  |       |            |   |  |  |
|-------------------|--|-------|------------|---|--|--|
| 1.                | Surname  |       | First Name |   | Middle Name                              |  |
| 2.                | Preferred mailing address for publications   |       |            |   |  |  |
| Permanent address |  |       |            |   |  |  |
| 3.                | Place  | Month | Day        | Year of Birth   | 4. Country of citizenship/naturalization |  |
| 5.                | Nature of work and title and/or military rank; name and address of organization with which you are associated. |       |            |   |  |  |
| 6.                | Check section or sections with which affiliation is desired.   |       |            |   |  |  |
|                   | <input type="checkbox"/> Geodesy   |       |            | <input type="checkbox"/> Oceanography                             |  |  |
|                   | <input type="checkbox"/> Seismology  |       |            | <input type="checkbox"/> Volcanology, Geochemistry, and Petrology |  |  |
|                   | <input type="checkbox"/> Meteorology   |       |            | <input type="checkbox"/> Hydrology                                |  |  |
|                   | <input type="checkbox"/> Geomagnetism and Aeronomy   |       |            | <input type="checkbox"/> Tectonophysics                           |  |  |

## 7. EXPERIENCE (List below, use added sheets as necessary)

|        |      |    |                                  |                               |
|--------|------|----|----------------------------------|-------------------------------|
| Dates: | From | To | Name and address of organization | Title, duties, nature of work |
|--------|------|----|----------------------------------|-------------------------------|

## 8. EDUCATION (List Below, use added sheets as necessary)

|        |      |    |        |         |               |                      |
|--------|------|----|--------|---------|---------------|----------------------|
| Dates: | From | To | School | Address | Major Subject | Degree, if any; year |
|--------|------|----|--------|---------|---------------|----------------------|

(over)



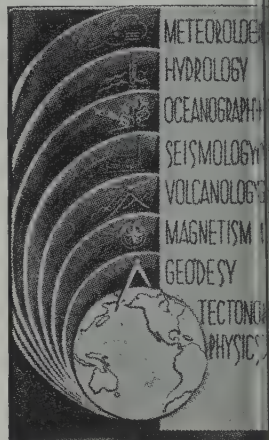
(Continued from previous page)

Those eligible as candidates for election to the grade of ASSOCIATE MEMBER shall be:

**ASSOCIATE MEMBER** Persons who have an active interest in physical processes of the Earth or technical assistance in the application of geophysics. In general, the minimum qualification for associate membership will be acceptable training or experience in some field of geophysics or allied science.

**CORPORATION MEMBER** Corporations and other interested organizations shall be eligible as candidates for election to CORPORATION MEMBERSHIP. They shall have the privilege of designating a representative who has the rights and privileges of Members (use special form).

**STUDENT MEMBER** Those eligible as candidates for election to the grade of STUDENT MEMBER shall be persons who are graduate or undergraduate students in residence at least half-time and who are specializing in the geophysical sciences. Teaching or research assistants enrolled in more than half of a full-time academic program may also be eligible for Student Membership. Student Members shall have all the privileges of Members except that they shall not vote or hold office.



-----  
Cut along this line

\*9. References: Please list below names and addresses of two or three references; include members of the AGU or others who know you well.

\*10. Titles of technical contributions or publications, particularly those in the geophysical sciences, and where published.

\*11. Brief statement of any special interests or qualifications in the geophysical sciences.

Date \_\_\_\_\_

\_\_\_\_\_  
*Written Signature*

12. (STUDENT MEMBERS ONLY) The person whose signature appears above is known to me and is a student majoring in \_\_\_\_\_ (subject) at \_\_\_\_\_

(Name of college or university) expected to graduate in \_\_\_\_\_ (year) with the degree of \_\_\_\_\_

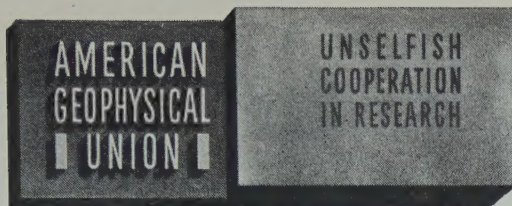
☐ He is a full-time student, or ☐ a teaching or research assistant enrolled in more than half of a full-time academic program.

\_\_\_\_\_  
(Signature of faculty sponsor) ☐ Check here if faculty sponsor is a member of AGU and willing to act as a regular sponsor for associate membership as well.

\_\_\_\_\_  
(Typed or printed name of sponsor)

\_\_\_\_\_  
(Title)

\* Applicants for student membership may omit Questions 9, 10, and 11, but must fill in Question 12. Please return form with check or money order payable to American Geophysical Union, 1515 Massachusetts Ave., N.W., Washington 5, D. C.



## INFORMATION CONCERNING CORPORATION MEMBERSHIP

The American Geophysical Union is a non-profit scientific organization established by the National Research Council. It is the American National Committee of the International Union of Geodesy and Geophysics, and its Executive Committee is the Committee on Geophysics of the National Research Council.

Extracts from the Statutes:

*Article 3. Membership*—The membership of the American Geophysical Union shall be as follows:

- (e) *Corporation Members*—Corporations and other organizations interested in geophysics elected by the Executive Committee of the Union. The designated representative of each such organization shall enjoy the privileges of a Member.

(Continued on next page)

-----  
Cut along this line

## American Geophysical Union

### PROPOSAL FOR CORPORATION MEMBERSHIP

To the Executive Committee, American Geophysical Union  
1515 Massachusetts Ave., N.W., Washington 5, D. C.

Gentlemen:

As an indication of our interest in the aims and activities of the American Geophysical Union, and to assist in maintaining and extending its program of publication and other work in the development of the geophysical sciences, the undersigned applies for Corporation Membership in the AGU and, until further notice, agrees to pay annual dues, currently at the rate of \$100 per unit of corporation membership, in accordance with the information set forth above.

Company or Organization \_\_\_\_\_

By \_\_\_\_\_ Title \_\_\_\_\_

(Signature)

(over)



(Continued from previous page)

Extracts from the By-Laws:

- (2) . . . Members of class (e) shall pay dues of not less than \$100 for each calendar year; . . .
- (21) One copy of each issue of (a) the *Transactions*, (b) *Journal of Geophysical Research*, (c) any published *List of Members and Officers*, and (d) any other publication which may be approved for *free distribution* to the membership by the Executive Committee of the Union, shall be sent to each . . . Corporation Member. . . Each . . . organization in good standing may purchase any available publication of the Union at a discount from printed price list to nonmembers. The General Secretary is authorized to establish discounts for sales of publications.

Action of the Executive Committee, November 29, 1946:

- (1) A list of corporation members shall be published on one or more pages immediately after the final page of text in each issue of the *Transactions*.
- (2) A list of corporation members shall be included in the Membership Directory as a distinct unit.

AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Ave., N.W.  
Washington 5, D. C.

-----  
Cut along this line

Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_

General fields of activity \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

The following person is designated as our representative in this membership \_\_\_\_\_  
Title \_\_\_\_\_

Number of units of membership desired (this will be taken as one unless otherwise indicated) \_\_\_\_\_

Place \_\_\_\_\_

Date \_\_\_\_\_

# Contents

(Continued from back cover)

|   | PAGE   |
|---|--|
| Region Traveling Disturbances and Sporadic-E Ionization . . . . .   | <i>L. H. Heisler and J. D. Whitehead</i> 2767                                      |
| Variable Atmospheric-Density Model from Satellite Accelerations . . . . .   | <i>Luigi G. Jacchia</i> 2775   |
| The Motion of a Satellite in an Asymmetrical Gravitational Field . . . . .  | <i>Peter Musen</i> 2783  |
| Further Radar Observations from an Earth Satellite . . . . .  | <i>J. E. Keigler and L. Krawitz</i> 2793   |
| The Nonseasonal Variations in Sea Level along the West Coast of North America   | <i>Gunnar I. Roden</i> 2809  |
| Latest Achievements of Physical Geodesy . . . . .   | <i>W. A. Heiskanen</i> 2827  |
| in Geodesy and Photographic Observations . . . . .  | <i>Alexander Corpacius</i> 2837  |
| Horizontal Movement in the Earth's Crust . . . . .  | <i>C. A. Whitten</i> 2839  |
| Simultaneous Adjustment of Angular and Distance Measurements . . . . .  | <i>Frederick F. Ceely, Jr.</i> 2845  |
| Bottom Reflection Measurements on the Continental Shelf, Bermuda Banks, West Indies Arc, and in the West Atlantic Basins    | <i>John Ewing, Bernard Luskin, Archie Roberts, and Julius Hirshman</i> 2849        |
| The Mass Transport through the Drake Passage . . . . .  | <i>Feodor Ostapoff</i> 2861  |
| Distribution Pattern of Rainfall in the Leeward Koolau Mountains, Oahu, Hawaii . . . . .                                    | <i>John F. Mink</i> 2869   |
| Evaporation Losses from Small-Orifice Rain Gages . . . . .  | <i>Harold E. Gill</i> 2877   |
| Isolation Measurements Based on Heat Flow Through Soil with Special Reference to Paddy Fields                               | <i>Seitaro Suzuki</i> 2883   |
| Fusion Effects in Miscible Displacement Occurring in Saturated and Unsaturated Porous Materials                             | <i>J. W. Biggar and D. R. Nielsen</i> 2887   |
| Radioactivity Associated with Underground Nuclear Explosions . . . . .  | <i>Roger E. Batzel</i> 2897  |
| Natural Radiocarbon in the Atlantic Ocean   | <i>Wallace S. Broecker, Robert Gerard, Maurice Ewing, and Bruce C. Heezen</i> 2903 |
| Volume Diffusion as a Mechanism for Discordant Lead Ages . . . . .  | <i>G. R. Tilton</i> 2933   |
| The Distribution of Cosmic-Ray-Produced Rare Gases in Iron Meteorites . . . . .   | <i>P. Signer and A. O. Nier</i> 2947   |
| The Origin of Meteorites . . . . .  | <i>Brian Mason</i> 2965  |
| Black, Magnetic Spherules in Sediments . . . . .  | <i>W. D. Crozier</i> 2971  |
| Principles of Structural Stability of Underground Salt Cavities . . . . .   | <i>Shosei Serata and Earnest F. Gloyna</i> 2979                                    |
| Geomagnetic and Solar Data . . . . .  | <i>J. Virginia Lincoln</i> 2989  |
| Letters to the Editor:  |  |
| Angular Motion of the Spin Axis of the Tiros I Meteorological Satellite Due to Magnetic and Gravitational Torques . . . . . | <i>William R. Bandeen and Warren P. Manger</i> 2992                                |
| A Method for Measuring Temperature Directly in the Upper Atmosphere with a Rocket-Borne Magnetic Mass Spectrometer          | <i>C. Y. Johnson, J. H. Hoffman, J. M. Young, and J. C. Holmes</i> 2996            |
| Outer Radiation Belt and Solar Proton Observations with Explorer VII during March-April 1960                                | <i>James A. Van Allen and Wei Ching Lin</i> 2998                                   |
| Solar Cosmic Rays and Soft Radiation Observed at 5,000,000 Kilometers from Earth  | <i>R. L. Arnoldy, R. A. Hoffman, and J. R. Winckler</i> 3004                       |
| The Effect of the Initial Phase of a Magnetic Storm upon the Outer Van Allen Belt   | <i>J. M. Malville</i> 3008   |
| Increase of the Nucleonic Intensity on May 4, 1960  | <i>J. H. Trainor, M. A. Shea, and J. A. Lockwood</i> 3011                          |
| A Search for Geomagnetic Singular Days . . . . .  | <i>T. Pohrte, C. Warwick, and N. Macdonald</i> 3013                                |
| Measurements of the Geomagnetic Field near Capetown . . . . .   | <i>J. R. Heirtzler and J. Hirshman</i> 3016  |
| A Proposed X-Ray Telescope for the 1- to 100-A Region . . . . .   | <i>Albert V. Baez</i> 3019   |
| Author's Reply to the Discussion by M. A. Benson of the Paper, 'A New Approach to Peak Flow Estimation' . . . . .           | <i>R. Rangarajan</i> 3021  |
| Electrical Resistivity of Frozen Earth . . . . .  | <i>T. Hatherton</i> 3023   |



# Contents

## SYMPOSIUM ON THE EXOSPHERE AND UPPER *F* REGION

(Edited by C. O. Hines)

|   |   |
|---|---|
| Summary of the Proceedings.....   | C. O. Hines   |
| The Exosphere and Upper <i>F</i> Region.....  | Francis S. Johnson  |
| Structure of the Earth's Exosphere.....   | S. F. Singer  |
| Whistler Dispersion and Exospheric Hydrogen Ions.....   | R. E. Barrington and T. Nishizaki                               |
| Electron Densities to 5 Earth Radii Deduced from Nose Whistlers.....  | R. L. Smith and R. A. Helliwell                                 |
| Radio Propagation Measurements Using the Explorer VI Satellite.....   | Carl D. Graves  |
| A Sounding Rocket Measurement of Electron Densities to 1500 Kilometers.....   | Warren W. Berning   |
| Comment on Models of the Ionosphere above $h_{\max}F_2$ .....   | J. W. Wright  |
| Electron-Density Distribution in the Upper Ionosphere from Rocket Measurements.....   | John S. Nisbet  |
| The Determination of Ionospheric Electron Content from Satellite Doppler Measurements. 1. Method of Analysis.....           | W. J. Ross  |
| The Determination of Ionospheric Electron Content from Satellite Doppler Measurements. 2. Experimental Results.....         | W. J. Ross  |
| Faraday Rotation Observations of the Electron Content of the Exosphere.....   | R. B. Dyce  |
| Radar-Lunar Investigations at a Low Geomagnetic Latitude.....   | George H. Millman, Armand E. Sanders, and Robert A. Mather      |
| Use of the Incoherent Scatter Technique to Obtain Ionospheric Temperature.....  | T. E. VanZandt and K. L. Bowles                                 |
| Some Characteristics of Ionospheric Backscatter Observed at 440 Mc/s.....   | V. C. Pineo, L. G. Kraft, and H. W. Briscoe                     |
| Radio-Wave Scattering by an Ionized Gas in Thermal Equilibrium.....   | J. A. Fejer   |
| Low-Energy Solar Cosmic Rays and the Geomagnetic Storm of May 12, 1959.....   | J. R. Winckler and P. D. Bhavsar                                |
| Observations of Solar Cosmic Rays Near the North Magnetic Pole.....   | K. A. Anderson and D. C. Enemark                                |
| Comparison of Solar Cosmic Ray Injections Including July 17, 1959 and May 4, 1960.....                                      | K. G. McCracken and R. A. R. Palmeira                           |
| Observations of Solar Flare Radiation and Modulation Effects at Balloon Altitudes, July 1959.....                           | A. Ehmert, H. Erbe, G. Pfozter, C. D. Anger, and R. R. Brown    |
| Development of Multiple Radiation Zones on October 18, 1959.....  | B. J. O'Brien and George H. Ludwig                              |
| Calculations of Cosmic-Ray Trajectories near the Equator.....   | Paul J. Kellogg   |
| Electrons of the Van Allen Radiation.....   | Paul J. Kellogg   |
| Geomagnetic Storm Sudden-Commencement Rise Times.....   | A. J. Dessler, W. E. Francis, and E. N. Parker                  |
| A Note on Harmonic Analysis of the Geophysical Data with Special Reference to the Analysis of Geomagnetic Storms.....       | Masahisa Sugiura  |
| Direct Measurement of Particles Producing Visible Auroras.....  | Carl E. McIlwain  |
| Audio-Frequency Electromagnetic Radiation in the Auroral Zone.....  | Georg Gustafsson, Alv Egeland, and Jules Aarons                 |
| Correlation of an Auroral Arc and a Subvisible Monochromatic 6300 Å Arc with Outer-Zone Radiation on November 28, 1959..... | B. J. O'Brien, J. A. Van Allen, F. E. Roach, and C. W. Gartlein |

(Continued inside back cover)